NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

EVALUATION OF CUMMULATIVE SUM (CUSUM) AND EXPONENTIALLY WEIGHTED MOVING AVERAGE (EWMA) CONTROL CHARTS TO DETECT CHANGES IN UNDERLYING DEMAND TRENDS OF NAVAL AVIATION SPARES

by

Les Wetherington

September 2010

Thesis Advisor: David Olwell
Second Reader: Ron Carlson

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2010	Master's Thesis

4. TITLE AND SUBTITLE
Evaluation of CUSUM and EWMA Control Charts to Detect Changes in Underlying Demand Trends of Naval Aviation Spares

5. FUNDING NUMBERS

6. AUTHOR(S) Les Wetherington

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE A

13. ABSTRACT (maximum 200 words)

The Navy must keep aircraft in a high state of readiness around the globe requiring spare parts to be available when and where needed. Managers need to know when changes in demand patterns are occurring far enough in advance to ensure continued availability of needed spare parts. This thesis presents an evaluation of two techniques using widely available software operating in a Windows environment to determine if changes are occurring in underlying demand patterns. These techniques are Cumulative Sum Control Charting and Exponentially Weighted Moving Average Control Charting. The use of the techniques was validated using a computer generated data set with known variation characteristics, and related processes were developed. After validation, the techniques were applied to four actual data sets with demand information from Navy aircraft. Both techniques proved effective with Cumulative Sum Charting providing slightly earlier alarms, and Exponentially Weighted Moving Averages being easier to use. Use of these techniques could allow detection of changes in time to mitigate the negative effects of the change and could be applied to a very wide range of processes. For the Navy, the widespread use of these techniques could lead to more aircraft being available for combat missions.

14. SUBJECT TERMS Cumulative Sum Control Charts, Exponentially Weighted Moving Average Control Charts, Changes in Spares Requirements, , Detection of Changes in Underlying Demand Trends

15. NUMBER OF PAGES 77

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

EVALUATION OF CUSUM AND EWMA CONTROL CHARTS TO DETECT CHANGES IN UNDERLYING DEMAND TRENDS OF NAVAL AVIATION SPARES

Les O. Wetherington, Jr.
Civilian, United States Navy, Patuxent River, Maryland
B.S., N.C. State University, 1977

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
September 2010

Author: Les Wetherington, Jr.

Approved by: Dr. David Olwell
Thesis Advisor

Professor of Practice Ron Carlson
Second Reader

Dr. Clifford Whitcomb
Chairman, Graduate School of Engineering and Applied Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The Navy must keep aircraft in a high state of readiness around the globe requiring spare parts to be available when and where needed. Managers need to know when changes in demand patterns are occurring far enough in advance to ensure continued availability of needed spare parts. This thesis presents an evaluation of two techniques using widely available software operating in a Windows environment to determine if changes are occurring in underlying demand patterns. These techniques are Cumulative Sum Control Charting and Exponentially Weighted Moving Average Control Charting. The use of the techniques was validated using a computer generated data set with known variation characteristics, and related processes were developed. After validation, the techniques were applied to four actual data sets with demand information from Navy aircraft. Both techniques proved effective with Cumulative Sum Charting providing slightly earlier alarms, and Exponentially Weighted Moving Averages being easier to use. Use of these techniques could allow detection of changes in time to mitigate the negative effects of the change and could be applied to a very wide range of processes. For the Navy, the widespread use of these techniques could lead to more aircraft being available for combat missions.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Maintaining aircraft in a high state of readiness around the globe is very challenging for the Navy. To keep the aircraft repaired and ready for missions, managers must allocate scarce resources to keep spare parts available for use by maintainers when and where needed. Responsible inventory managers must make predictions for future demand using historical demand data. This works well if the future demand doesn't deviate too much or too quickly from the historical demand. However, if the underlying data distributions of the spare requirements change, the predictions made can be in error leading to too many spare parts if demand is falling or a lack of spare parts if demand is increasing. Too many supply parts indicate that resources were not used efficiently. Too few spare parts can lead to aircraft not being available for their missions due to lack of the correct repair parts. Thus it is very important for managers to know if the underlying demand distribution of the spares requirements is changing.

While managers want to know quickly about a change, they do not want false alarms. The tradeoff between fast response to changes and the number of false alarms is a key design consideration for any monitoring system and is discussed in detail in this thesis.

In order to monitor any process, it is very important to know what data is available and representative of the process. Within the Navy, a component is coded as Beyond Capability of Maintenance (BCM) if it cannot be repaired by the fleet and a requisition for a replacement component is prepared. One could use either the requisition data or the BCM data as representative of the total demand on the wholesale supply system. Historical BCM data is much easier to obtain than historical wholesale demand data, thus BCMs were chosen as the data element to be analyzed.

One goal of the thesis was to develop the processes to create CUSUM and EWMA charts using widely available tools. Thus Minitab 16 was chosen as the statistics program as it is used widely in industry and the military. Another program, *Anygeth.exe*,

is needed to determine the correct variables for CUSUM charts. This program is available for free on the internet and directions are contained in the thesis to download the program.

The research was performed in two phases: validation of methodologies, and analysis of components. To first validate the analysis methodologies, a data set with known characteristics was analyzed using CUSUM and EWMA control charts. The data set was formed by combining two data sets with the first 500 data points following a Poisson distribution with a mean of 0.5, and the second 500 data points following a Poisson distribution with a mean of 0.7. This computer generated data set was created to follow the type of variation, a persistent shift in mean, for which a CUSUM control chart is most exactly tailored to detect. By comparing CUSUM and EWMA charts generated against a data set with known characteristics, the methodologies were validated to develop the CUSUM and EWMA charts using *Anygeth.exe* and Minitab.

Next, CUSUM and EWMA control charts were generated for the remaining four real data sets and the results were compared to each other relative to the known variability in the data sets. The effectiveness of the charts relative to determining shifts in underlying trends and the efficiency of the charts relative to the time and expertise required to generate the charts were compared.

Both CUSUM and EWMA were capable of detecting shifts in demand data. CUSUM generally provided a slightly faster alarm, but required considerably more expertise and time to use. A unique set up along with an understanding the underlying data distribution was required for each component when developing CUSUM charts. EWMA charts tended to be slightly slower in providing alarms, but were much easier to set up with a single set up required for all components for a given ARL. Overall EWMA charts are more efficient to use with a slight loss in effectiveness. With the use of these tools, Navy managers could take a more proactive response to issues enabling more aircraft to be in a state of combat readiness.

While this thesis focused on a specific issue of spares for Navy aircraft, the concepts and methodologies developed within this thesis would readily apply to any

process for which the user wanted to detect changes in sufficient time to allow mitigating actions. Recommendations were made for further research that could automate these processes and provide more information about how to solve the specific issue in addition to detecting the issue.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

3M	Maintenance, Material, and Management
ARL	Average Run Length
BCM	Beyond Capability of Maintenance
CL	Center Line
CUSUM	Cumulative Sum
DI	Decision Interval
EWMA	Exponentially Weighted Moving Average
FH	Flight Hour
LCL	Lower Control Limit
MAF	Maintenance Action Form
NAVICP	Naval Inventory Control Point
RFT	Ready for Tasking
SPC	Statistical Process Control
SRA	Shop Replaceable Assembly
UCL	Upper Control Limit
WRA	Weapons Replaceable Assembly

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my leadership at Naval Air Systems Command for selecting me for this program and for their continuous support and patience while I was going through the voyage. This includes Mr. Andy Monje, Mr. Brian Gilp, Mr. Craig Paylor, Mr. Stu Young, Dr. John Fischer, and Professor of Practice Ron Carlson. Professor of Practice Carlson also served as my second reader, and for that I am very appreciative.

I also owe thanks to Dr. David Olwell for serving as my primary advisor and for his patience, wisdom and expert guidance as I worked through this unfamiliar territory and also for his assistance during the past several years of my career.

Additionally, I owe thanks to Mr. Larry Garvey, Mr. Larry Croll, and Ms. Judy Poole. For many years they have helped to educate me on the complexity and intricacies of the Navy supply system and they continued to do so during the research on this paper.

I would like to thank Mr. John Hargraves, Mr. John Martin, and Ms. Apryl Gilbert for assistance with the data for this thesis and helping me understand how all the thousands of pieces in the system tie together.

I owe thanks to Mr. Kevin Wiersma who provided the core idea for this thesis and continued to help with information and expert knowledge as this thesis progressed.

I thank Ms. Pam Silva and Professor Mary Vizzini for your assistance in reviewing this thesis for structure, language, and logical organization. It is a much better paper due to your efforts.

I also thank Rosalie who somehow found the patience to grow a relationship with me when I had almost no time or energy to give. That speaks volumes for you.

I thank my family for their amazing support and patience while I juggled the requirements of school, work and family. To my Mom, my sisters Ilene and Anna, and my children Kasi and Evan, I will always remember how you stood by me. Without your support and understanding, I would not have accomplished everything I did.

And finally, I owe special thanks to my Dad, to whom I dedicate this thesis. In your last days from your hospital bed, you remained curious about my school work and continued to encourage me. I thank you for your gifts of curiosity and wonder.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

U.S. Naval Air Forces face the difficult challenge of supporting aircraft in harm's way in locations covering the globe. Behind the scenes, the Navy manages a massive repair supply system with the complex task of keeping the right repair parts in stock at the right place and right time with limited resources and with constantly changing mission requirements. Inventory managers must make predictions about future demand for parts in order to allocate scarce resources in sufficient time for parts to be available when needed. With limitations on existing prediction techniques and funding, the components needed to return an aircraft to Ready For Tasking (RFT) status are many times not available when and where they are needed, leading to specific aircraft being placed in non-RFT status and thus unavailable to perform the required missions.

The Navy's supply system currently relies on data taken from requisitions for parts from military customers to make predictions of future supply requirements, and bases the planned repair rate in commercial and organic depots on these predictions. This data is called "demand" data since it is based on the demand for parts by users. The demand data is often characterized by spikes when there is a sudden increase in demand and plateaus when demand falls to zero for periods of time.

Often, it is very difficult for a manager to predict future demand, as the manager cannot tell whether a spike or plateau in the data represents a true change in the underlying demand distribution, or just represents a transient signal. Reacting to a transient signal or false alarm, and not reacting to a true demand shift signal both have negative consequences, although they may not be symmetrical with one another. Reacting to a false alarm causes the needless expenditure of scarce resources and the diversion of those resources from other areas where they may be critically needed. Not reacting to a true increasing demand signal can cause aircraft to be unavailable for critical

missions due to lack of spare parts. Thus, managers need to be able to differentiate between a true demand change and a transient signal in order to most effectively support Navy aircraft.

By using the proper statistical methods, a manager can design a system to detect true demand distribution changes with a known false alarm rate, thus limiting false alarms to an acceptable level. Reducing false alarms does come at a price and that price is the delay in accumulating enough evidence to generate a signal. However, the user may consciously balance the false alarm rate while in control with the reaction time delay when out of control to design the alarms based on the criticality or price of the component under study. If the responsible managers can receive valid warnings of changes in usage trends in sufficient time, they can take more informed mitigating actions. For example, a manager could increase production or correct the cause of the change in trend before a zero shelf stock condition occurs and aircraft becomes non-RFT for lack of repair parts. Thus, the use of a proper statistical method serves as an effective tool in the life-cycle management of aviation assets. Managers may use these tools to optimize the availability of Navy aircraft in the fiscally restrained repair supply system allowing our naval aviators to do their important missions for our nation's defense.

B. PURPOSE OF STUDY

This thesis will demonstrate and compare the effectiveness of two statistical methods that may be useful to meet the manager's requirement for detecting changes in underlying demand for spare components with known false alarm rates: CUSUM and EWMA control charts. In this case, they will be applied to the BCM actions from the entire fleet, per week as this data is much more widely available within the Naval Aviation community than the related supply demand data. Fleet BCMs and supply demand are highly correlated and are discussed in detail in Chapter II.

C. RESEARCH QUESTIONS

1. What demand representative data is collected within U.S. Naval Air Forces for which current and historical data is readily available?

2. How effective are CUSUM and EWMA control charts in detecting changes in underlying demand patterns with a pre-determined false alarm rate?

3. What processes and level of expertise are required to use these techniques within the Naval Aviation environment?

4. Does one technique stand out as being more efficient or more effective and thus should be recommended for use in this environment?

D. BENEFITS OF STUDY

The evidence analyzed in this paper reveals that EWMA control charts are nearly as effective as CUSUM charts while requiring much less user skill and expertise to implement. Thus, more managers could use this technique to monitor the Navy maintenance system, and allow us to keep more aircraft in RFT status using fewer scarce resources. As discussed in the following paragraph, these findings may have general application well beyond aircraft maintenance.

E. SCOPE AND METHODOLOGY

The scope of this paper covers the use of CUSUM and EWMA techniques with commercially available software to detect changes in demand patterns for naval aviation spares. However, the findings may be equally as applicable to any type of process under statistical control for which the reader desires to know if the underlying distribution of data is changing.

For the overall methodology, the author validated the process of developing CUSUM and EWMA charts using a designed data set, compared the results from CUSUM and EWMA analysis of four real data sets, and provided a generalized analysis of the efficiency and effectiveness of the techniques. For both the CUSUM and EWMA techniques, the Average Run Length (ARL), or the time between false alarms, was set at approximately 100 weeks. The validation data set was for an imaginary component "Z", with known distribution and change in distribution, such that the change in distribution was designed specifically to optimize detection by CUSUM. The four real data sets contained the demand figures for actual naval components in operation in the fleet where

the variability in demand ranged from very low to extreme. By using four data sets spanning the range of variability as seen in the fleet, the analysis could be applicable to a wide range of systems in operation in the Navy.

Minitab version 16 and Excel 2007 for Windows were used in the analysis, as well as a program called *Anygeth.exe*, that is available for free via the Internet. Directions to obtain *Anygeth.exe* are provided in a later section of this paper. All of these programs are available to most users who might wish to explore the techniques used in this paper.

In summary, the methodology consisted of examining the effectiveness and efficiency of the CUSUM and EWMA techniques in regards to their ability to detect changes in underlying data distributions of representative data sets, while maintaining a predetermined false alarm rate. The findings could be of value to any reader who desires to monitor any process.

II. AIRCRAFT MAINTENANCE OVERVIEW

A. INTRODUCTION

This section will provide a general overview of the operations of aircraft within the Naval Air Forces and discuss how maintenance is performed and data is collected during maintenance and operations. While a massive amount of data is gathered by the Navy, it is important to choose the right data to analyze to detect specific changes. By showing how the data is gathered and what the data represents, the author will provide the rationale for the use of analysis of BCMs as the main indicator of changes in the supply demand patterns for the aircraft. With sufficient warning time of a change in demand pattern, Navy managers may be better able to take actions that would reduce the negative impacts of the change, such as running out of inventory of spare parts, and thus be able to keep aircraft availability high.

B. THREE LEVELS OF AIRCRAFT MAINTENANCE

The Navy has three levels of maintenance. They are: the Organizational Level or “O level,” Intermediate Level or “I level,” and the Depot Level or “D Level.” At the O level, the primary focus is on returning an aircraft to RFT status. This includes routine servicing of the aircraft, and removal and replacement of failed components or Weapons Replaceable Assemblies (WRAs) as they are called within the Navy. At the I level, the primary focus is on repairing failed WRAs that have been removed by the O level. Typically, WRAs are repaired by exchanging modules or Shop Replaceable Assemblies (SRAs) at the I level. Both O and I levels are mostly staffed with military personnel. These military personnel may deploy to locations all over the globe as required by operational demands of the services. The D level is staffed primarily with civilians and these civilians do the most extensive maintenance of all levels. WRAs and SRAs that are declared Beyond the Capability of Maintenance (BCM) from the I level are repaired at the D level.

C. FLOW OF DATA AND COMPONENTS WITHIN THE NAVAL AIR FORCES

The tracking of maintenance data begins when a component fails on the aircraft. The failed component is removed at the O level and the data trail begins. Thus, when the O level determines that a WRA has failed, the WRA is removed from the aircraft and this removal is recorded on a Maintenance Action Form (MAF) within the 3M system, a data system used by the Navy and Marine Corps where the “Ms” stand for “Maintenance”, “Material”, and “Management”. The suspected to be faulty WRA is taken to the local supply organization and exchanged for an “A” condition WRA. (“A” condition means that the component fully functional and is ready for installation on an aircraft.) The “A” condition WRA comes from retail stock, if stocked locally, or wholesale stock, if stocked remotely from a centralized location. A local retail stock is in place if there is a history of demand and readiness impact shows that it is cost effective to stock the items locally. The organizational level maintenance activity will then install the “A” condition WRA into the aircraft, run any associated Built In Test, record the installation on the MAF, and ready the aircraft for its next flight, returning the aircraft to RFT status. This completes the O level action for this failure.

The suspected failed WRA, on the other hand, now begins its path through the system. It is now considered “F” condition stock, meaning that it is not functional. The local retail supply will send the WRA to the appropriate I level maintenance activity for repair. If the I level maintenance activity can repair the WRA, they do so and return it to the local retail supply as an “A” condition asset (replacing the previously issued “A” condition asset). The I level maintenance activity will indicate on a MAF in the 3M system that this was verified as a failure and that they repaired the item. If the I level maintenance activity cannot repair the item, then they will code it as a BCM in the 3M system and return it to the local retail supply organization as “F” condition stock. The local retail supply organization will then return the “F” condition stock to the wholesale supply level and a supply document known as a replenishment requisition will be submitted for a replacement item for local retail stock. From the wholesale level, the “F” condition asset will be scheduled for induction into a D Level maintenance activity for

repair as scheduled to maintain the “A” condition asset level with control limits at the wholesale level. The wholesale level may also purchase new units to replace those that cannot be economically repaired at the D level. Through the processes discussed in this paragraph, the failed WRA is repaired and returned to the supply system for use, or scrapped and a new WRA is procured to replace it.

What is important to note here is that demand patterns may be very different at the different levels of the system with very different impacts on aircraft availability. If the I level has significant repair capability, there could be large number of failures on the aircraft and subsequent demand upon the retail level stock, but little or no demand upon the wholesale level. As long as the I level can repair components and the O level can replace them on the aircraft in a timely manner, the availability of the aircraft will remain high regardless of the failure rate of the component. (The cost of maintaining highly unreliable components is another issue and is not considered within this paper.) However, if the I level loses the equipment or skilled personnel to do a repair, the demand at the wholesale level could easily spike. For example, if the I level has full capability to repair a generator from an aircraft, the wholesale level may see no demand at all for the generator as long as the I level can repair all of generators that fail. However, if the I level test bench for the generator fails and cannot be repaired, all of the generators that fail on the aircraft will be coded as a BCM from the I level and become demand upon the wholesale level. The failure rate on the aircraft may not have changed, but the wholesale demand rate can spike up suddenly due to lack of repair capability at the I level. It is this unanticipated spike in wholesale demand, above the level of normal demand, that often leads to aircraft being unavailable for use due to lack of “A” condition parts. For this reason, the author chose wholesale demand as the variable to analyze using CUSUM and EWMA.

While the 3M system records BCMs, the supply system operates primarily on requisition documents as a measure of wholesale demand. The timing of BCMs and supply requisitions was examined and it was found that with few exceptions, they are very close in timing. Thus, one may use either the wholesale requisition data or the BCM data as representative of the total demand on the wholesale supply system. Since the

author was primarily concerned with wholesale demand in this analysis, and historical BCM data is much easier to obtain than historical wholesale demand data, BCMs were chosen as the data element to be analyzed.

D. CHAPTER SUMMARY

Maintenance of Naval Aircraft occurs around the globe in a complex three-level maintenance system. Data is recorded tracking the maintenance steps involved from removal of a component from the aircraft all the way through its repair and reinstallation. Data available in the fleet 3M system for BCMs is closely mirrored by requisition data available in the supply system and is representative of wholesale demand. When wholesale demand spikes and the supply system cannot provide sufficient repaired parts, aircraft availability can fall rapidly. Thus BCM data will be used in this paper for analysis of the research questions posed by this thesis. The author will analyze the use of CUSUM and EWMA techniques to detect changes in underlying demand distributions

III. RELATED RESEARCH

A. INTRODUCTION

Demand for spare parts in aviation is a complex and often studied field. When one examines research on the analysis and prediction of usage rates of spare parts in aviation, what becomes clear is that there is no consensus on a best method. While Varghese and Rossetti (2008) reported "no statistical difference" in techniques, Ghobbar and Friend (2003), determined that weighted moving averages were superior. Other researchers reached yet different conclusions. Clearly it is hard to predict demand for aviation spares and important that demand trends be monitored for change. The author will discuss the results of the literature review and the recommendations of a number of researchers in the field in more depth in the following paragraphs with rationale for the direction and focus of this thesis.

B. SUMMARY OF LITERATURE REVIEW

A number of papers discussing the issue of supply support in aviation and similar systems were reviewed. It is a very complex field and while it has been often studied, there appears to be little consensus on the best methodology to use. Demand in the aviation field is often not steady or easily predictable. This may be caused by deployment cycles, high time cycles where items are removed before failure when they reach a certain age, changes in operational environment or tempo, or a number of other factors. There may be long periods of low or no demand for an item followed by periods of high demand. As discussed by Williams (1982) in "Reorder Levels for Lumpy Demand," classical methods may not work well due to demand being "either zero or a large lump." Croston (1972) determined that "demand for constant quantities at fixed intervals may generate stock levels of up to double the quantity really needed." Clearly, the demand for aviation spares can be highly variable and very difficult to predict.

Some researchers found no statistical differences in techniques used. Varghese and Rossetti (2008) in "A Parametric Bootstrapping Approach to Forecast Interrmittent Demand" evaluated five different forecasting techniques including Croston (0.1),

Croston (0.2), Syntetos (0.1), Syntetos (0.2) and MC ARTA. They found that "there is no statistical difference between the best forecasting techniques and the other forecasting techniques."

Yet other researchers did report "best" methods. Eaves and Kingsman (2004) determined that the "best forecasting method for a spare parts inventory is deemed to be the approximation method." Ghobbar and Friend (2003), in their paper "Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model," examined the results of 13 different forecasting methods. These methods included additive winter, multiplicative winter, seasonal regression model, component service life, weighted calculation of demand rates, Croston, single exponential smoothing, exponentially weighted moving average, trend adjusted exponential smoothing, weighted moving averages, double exponential smoothing, and adaptive response rate single exponential smoothing. Their research results "confirm the continued superiority of the weighted moving average…" Gardner (1985) in "Exponential Smoothing: The State of the Art," states that "the empirical evidence favours Holt's models for trends…" In discussions with managers at Naval Inventory Control Point (NAVICP) from Philadelphia, PA, the author was informed that moving averages were used most often to predict future demand within the naval air forces.

It is therefore very important to monitor demand for spares as there is great variability in the demand and it is hard to predict changes. Gardner (1985), discussed CUSUM as a methodology to monitor forecast errors. Hawkins and Olwell (1997) discussed in depth the use of CUSUM to monitor process changes. Neubauer (1997) evaluated EWMA methods to detect shifts in processes and provided a very limited comparison to CUSUM charts. While CUSUM and EWMA control chart tools are readily available in Minitab, the author found no evidence of a direct comparison of these tools to monitor changes in demand for aviation spares. With the high variability in demand patterns, lack of consistent predictive tools, and the criticality of spares availability for naval aviation, the author decided to compare and evaluate CUSUM and EWMA control charts as a method to detect changes in demand for aviation spares.

C. CHAPTER SUMMARY

Estimating demand for aviation spares is a complex and controversial field. The demand patterns are often extremely variable and the cost of the items can be very high. A large amount of research has been performed over many years, yet the conclusions are mixed and there is no consensus on a "best" method to determine changes in demand. Since the "best" method for forecasting may be situation dependent, there will always be a need to monitor demand data for changes in underlying demand distributions. Thus, this thesis will analyze the use of EWMA and CUSUM control charts in the detection of the underlying demand distribution changes.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. OVERVIEW OF CUSUM AND EWMA CHARTING

A. INTRODUCTION

In all processes, there is variation present and aircraft maintenance is no exception. For managers charged with ensuring that there are adequate spare parts to repair aircraft, it is important to know if changes in demand are just a part of normal variability in the maintenance, or if the changes represent something different, namely a new higher demand pattern that could potentially use all available spares before they could be replenished. Additionally, the manager is also concerned if the change represents a true lower demand pattern such that resources could be shifted away from repairing components with reduced demand and towards more critical items. As discussed in Chapter III, many different methodologies have been used to analyze and forecast the demand for aviation spares. Chapter III concluded with the objective to compare CUSUM control charts to EWMA control charts as a method to detect changes in underlying demand trends for aviation spares. Thus, in this chapter, the CUSUM and EWMA control chart techniques will be discussed including the theory and possible application to the determination of changes to the distribution of spares demand. The information on CUSUM charting contained within this chapter was largely obtained from the book "Cumulative Sum Charts and Charting for Quality Improvement" (Hawkins & Olwell, 1997). The information on EWMA charting contained within this chapter was largely obtained from the article "The EWMA control chart: properties and comparison with other quality-control procedures by computer simulation" (Neubauer, 1997).

B. UNDERLYING CUSUM THEORY

1. Sources of Variability

Sources of variation in processes can be categorized as "common and special". Special causes are those for which a source can be identified. Identifying and removing special causes improves the process quality. What is left after the special causes are all removed is the "purely random variability," which is called "common cause variability" (Hawkins & Olwell, 1997). Common cause variability is inherent in the nature of the

process even while it is in control and cannot be improved without fundamental process change. Since the author is concerned with processes that are changing, this paper will focus only on special variability and not common cause variability.

There are two types of special causes for variability: transient and persistent (Hawkins & Olwell, 1997). Transient causes appear for a period of time, but then they go away. They may appear at a later time, or not. Persistent causes remain until detected and corrected. These two types of special causes leave different evidence and, thus, are best detected using different methods. CUSUM is designed to detect the persistent changes even if very small.

2. Detecting Transient Special Variability

Shewhart Xbar and R charts are widely used to detect transient special cause variability. Figure 1 shows an example of an XBar chart. An Upper Control Limit (UCL) and Lower Control Limit (LCL) bound the Center Line (CL). The Center Line represents the true in control mean of the process and the UCL and LCL are positioned three standard errors above and below the CL, respectively.

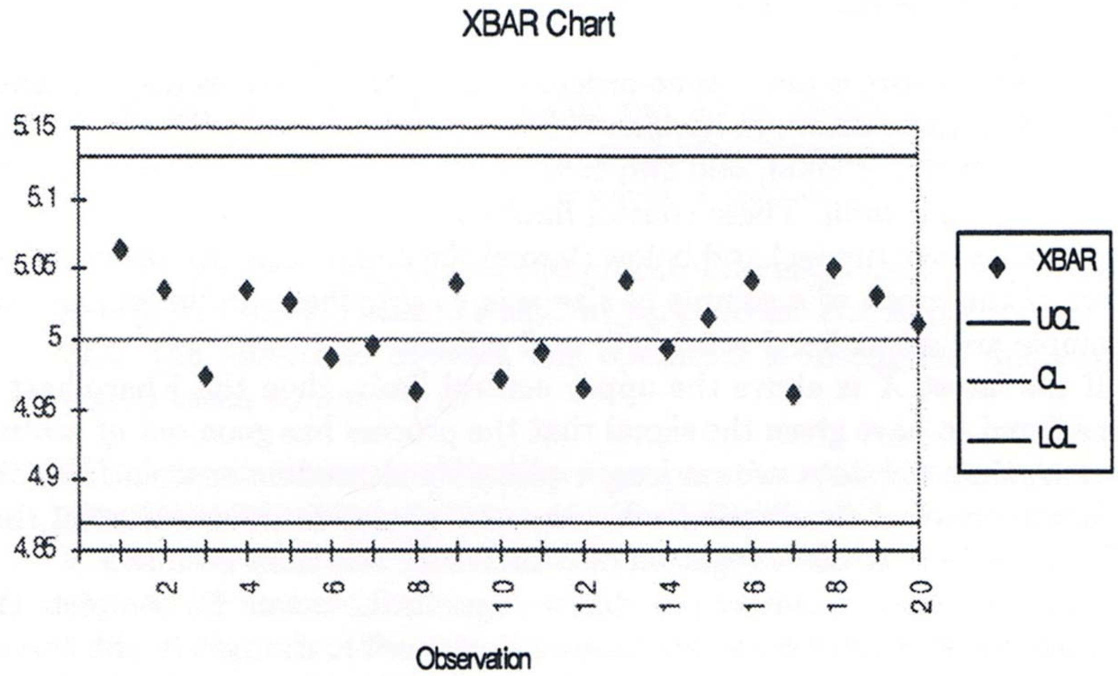

Figure 1. Shewhart Xbar chart from Hawkins and Olwell, 1997

The XBar chart monitors the mean and a similar chart called an R chart monitors the range of the sample readings. As long as the readings remain within the range between the LCL and UCL, the process is considered within control, as stated in Cumulative Sum Charts and Charting for Quality Improvement:

> The Shewart chart has a beautiful simplicity to it. It may also be as valuable for what it prevents as for what it motivates. As long as the points plot inside the control limits, no action is taken to alter the process. This rule can stop much unproductive tinkering that could take a process from a good state into a bad one. The control limits are placed sufficiently far from the center line that very few samples should plot outside them if the process remains in its in-control distribution. This means that when the control chart does give a signal, it should be taken very seriously. These attractions of the Shewart chart should not blind one to a serious limitation. It has no memory, and so although it is very effective for detecting isolated special causes that lead to large shifts in the data, it is not very effective in detecting more moderate shifts, *even if* these more moderate shifts persist. (Hawkins & Olwell, 1997)

The quote above discussing what control charting prevents is very important. In the author's thirty plus years of experience within Naval Aviation, many times scarce resources are diverted chasing down data point outliers or, as those within the business call them, "wild goose chases." Without the proper tools, managers many times don't know what is important and what should be left alone. Control charting allows a manager to know what is acceptable to ignore and allows focusing on the more critical issues.

Figure 2 shows a chart very similar to Figure 1, except the final ten readings have been increased by 0.03mm. It is very hard with the eye to detect that anything has changed in this chart, and there is no alarm or trigger.

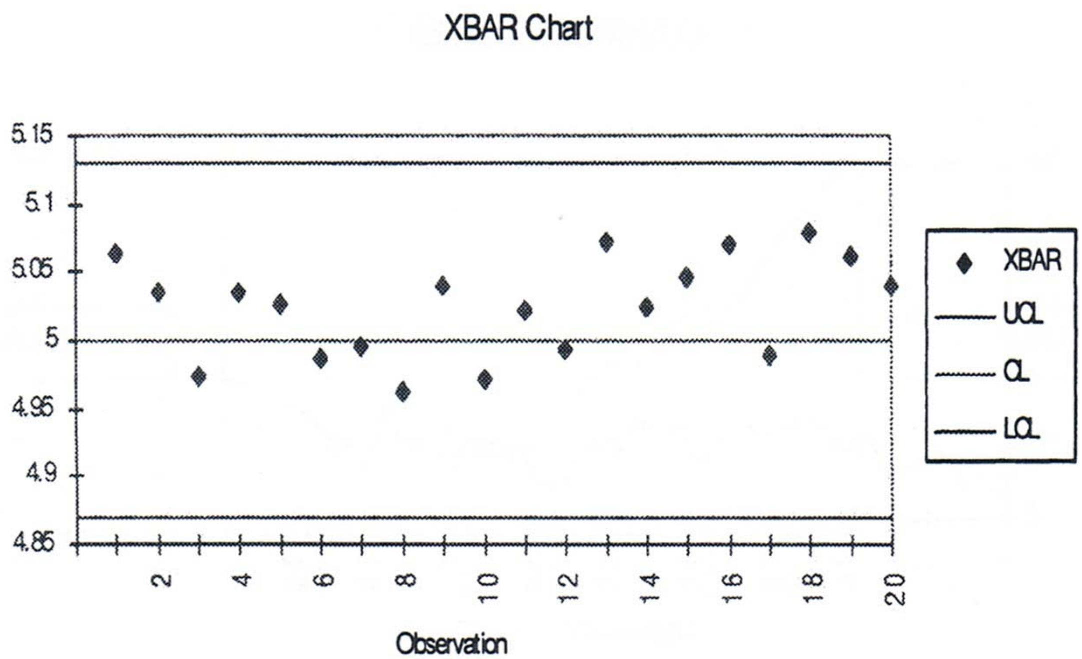

Figure 2. Shewhart Xbar chart with the last ten readings increased by 0.03 mm from Hawkins and Olwell, 1997

The reader can see that a Shewhart chart is not very effective in detecting small persistent changes in mean. Many attempts have been made to correct this problem, including adding "run rules" to the detection scheme. An example of a run rule is that "the process is out of control if two out of three successive points plot more than two

standard deviations from the center line.” (Hawkins & Olwell, 1997) A more effective method is to use CUSUM charting. One of the simpler explanations of CUSUM charting is contained in the document obtained from the Minitab corporate Web site “Using CUSUM Charts to Detect Small Process Shifts” (Bower) as follows:

> Though CUSUM charts have been well researched and developed, it is true that many quality practitioners do not use them, even though there may be justifiable reasons to make use of this technique for their process. Possibly this may be due to a lack of instruction on CUSUM charts in many classes on SPC. In practice, however, I find that many of these same quality practitioners play the game of golf; hence they are in fact already well versed in the technique behind CUSUM charts. In essence, for each hole in a round of golf, there are a specified number of times in which one should strike the ball, until it eventually drops into the hole. For example, on a par 4, if you strike the ball 4 times and it falls into the cup, then you held par. If you were able to do this task with only three shots (a “birdie”) then you are "1 under par" hence your cumulative sum is -1. This is continued throughout the course, the ultimate winner therefore having the lowest CUSUM. Picture a golfer who is holding par for the first 13 holes, then suddenly hits form and has five successive birdies towards the end of the round. The final CUSUM is therefore –5, though from viewing a CUSUM chart it would be clear to see when the “process” shifted.

In essence, the CUSUM is just the summation of each point in the process with the “par” or mean subtracted. While adding a stroke to a single hole may not make much of a difference to the golf game, persistently adding this stroke to each hole makes a large difference. This is precisely what CUSUM charting is designed to detect.

The simple CUSUM as discussed by Hawkins and Olwell (1997) is defined in Equation 1 as:

$$C_n = \sum_{j=1}^{n} (X_j - \mu)$$

Equation 1 - Basic CUSUM Equation

X_j is the value of the *j*th reading, μ is the in-control mean, and the difference between the two is summed from $j = 1$ to $j = n$. The value of C_n is then plotted against n and this forms the CUSUM chart.

If this value is plotted at each point in the data set for the values in Figure 1, the chart in Figure 3 is obtained.

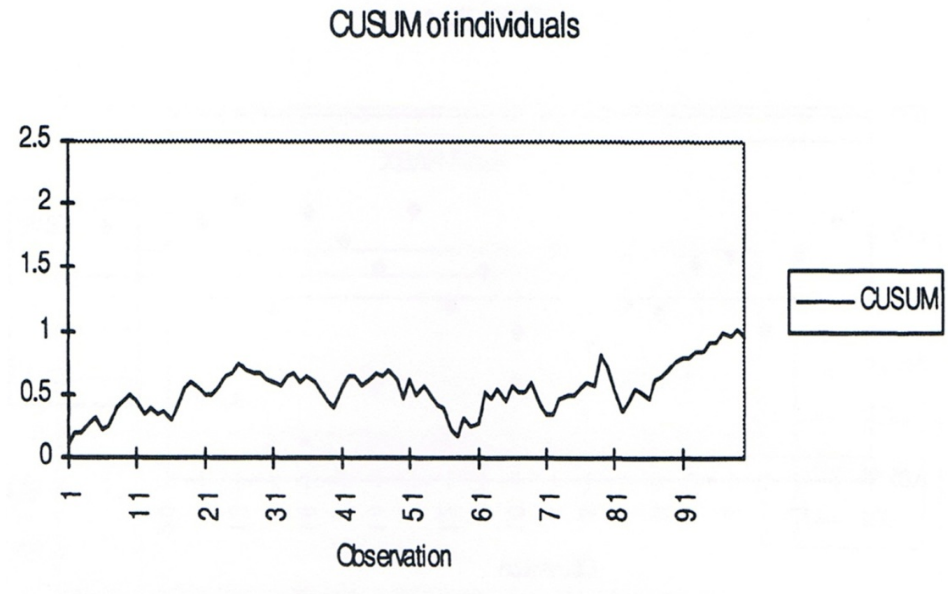

Figure 3. CUSUM of Original Diameters from Figure IV-1 from Hawkins and Olwell, 1997

Figure 3 shows a drifting randomness with no real trend in the data. From this chart, the reader could conclude that no shift in mean had occurred. In Figure 4, the data from Figure 2 with the small shift in mean is plotted on a CUSUM chart. It becomes very apparent on this chart that something significant has shifted around the sixtieth data point as the reader will note a steep shift of the CUSUM line towards the top right corner of the chart.

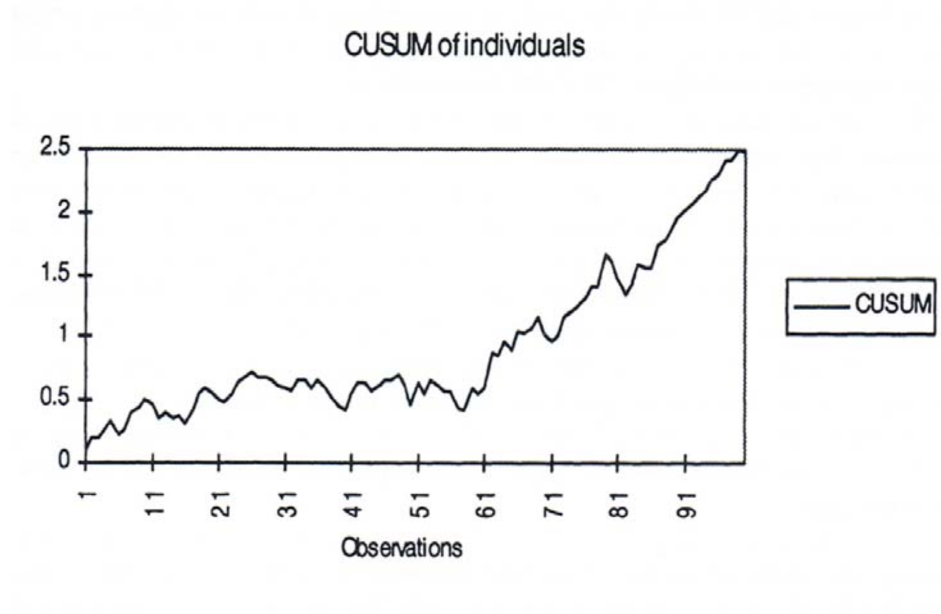

Figure 4. CUSUM of Shifted Diameters from Figure IV-2 from Hawkins and Olwell, 1997

In order to know when a CUSUM chart signals an out of control condition, a V-mask tool is sometimes used. This is called V-mask because of the shape of the mask as seen in Figure 5 where the process has been out of control since observation 70.

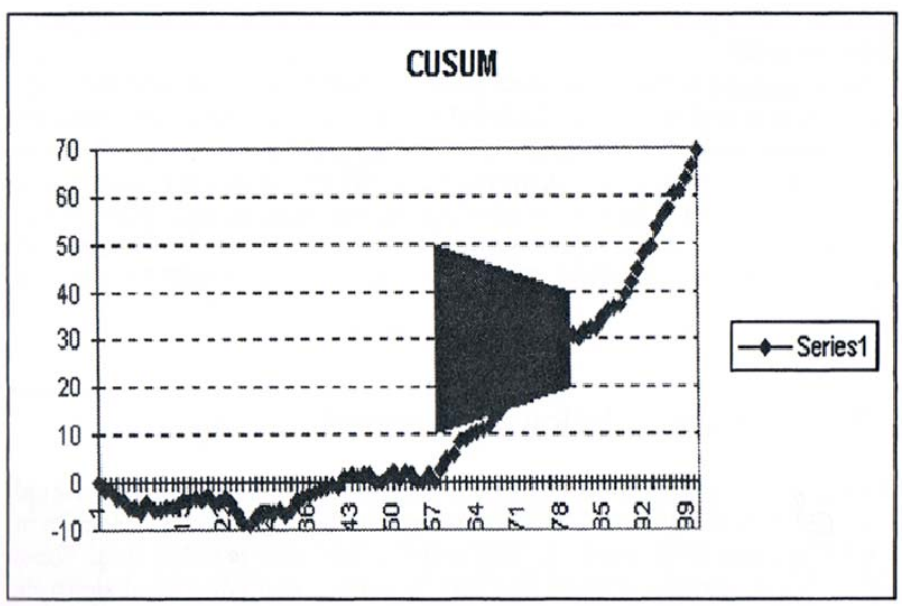

Figure 5. CUSUM V-Mask from Figure 1.8 from Hawkins and Olwell, 1997

The shape and size of the mask is derived from two parameters, *h* and *k*. As discussed by Hawkins and Olwell (1997):

> The mask has a flat front (right) end of height 2h and two legs of slope k and –k. As each new point is added to the CUSUM, align the center of the front edge with the point just plotted, and see whether all previous points are contained in the mask. If they are, then you conclude that the process is still in control – that its mean has not shifted. If any preceding point projects outside the mask, then you conclude that the mean has shifted.

The first parameter, *k*, tunes the CUSUM to react to a shift of a certain size. The second parameter, *h*, sets the average run length while in control.

A more concise form of the CUSUM chart is the Decision Interval (DI) form that is used within the Minitab software. This is equivalent to the V-mask version of the CUSUM with a slope *k* and leg height *h*. (For a complete discussion of out-of-control signals and the derivation of the associated mathematics, the reader is encouraged to consult Chapter 1 of Cumulative Sum Charts and Charting for Quality Improvement (Hawkins & Olwell, 1997). An example of a DI CUSUM chart is shown in Figure 6. An examination of this chart shows that the process clearly shifted around the sixtieth point.

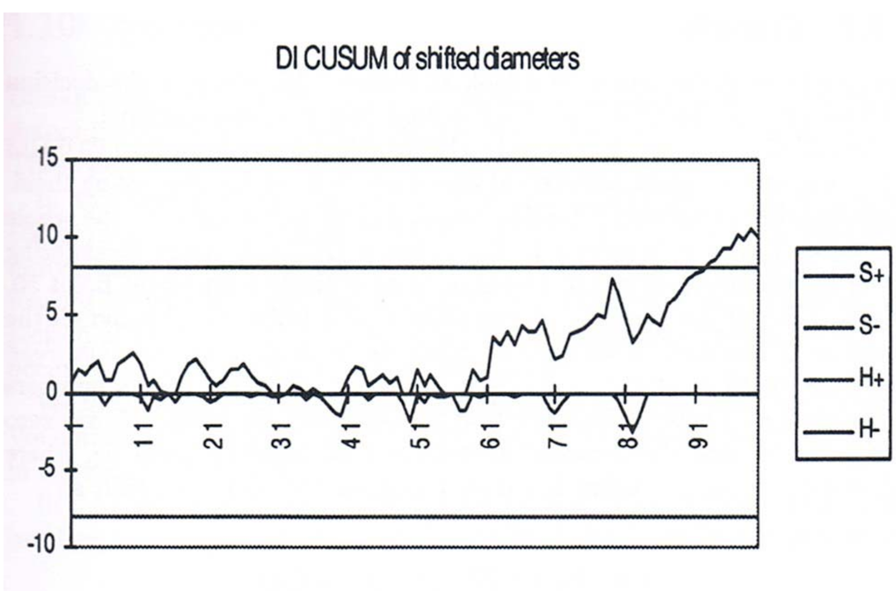

Figure 6. CUSUM Plot of Shifted Diameters from Figure 3 Hawkins and Olwell, 1997

The DI form of the CUSUM to test for an upward shift in mean is set up as follows as discussed by Hawkins and Olwell (1997):

$$C_0^+ = 0$$
$$C_n^+ = \max(0, C_{n-1}^+ + X_n - \mu - k)$$

This produces an alarm showing an upward shift in mean if $C_n^+ \succ h$. Similarly, the DI form of the CUSUM to test for a downward shift in mean is set up as follows:

$$C_0^- = 0$$
$$C_n^- = \min(0, C_{n-1}^- + X_n - \mu + k)$$

This produces an alarm showing an upward shift in mean if $C_n^- \prec -h$.

Determining the correct values of *h* and *k* to use is essential in creating a CUSUM system that will set off an alarm quickly when the mean has shifted without excessive false alarms during in control operations. The average time between alarms is termed the Average Run Length (ARL). Two ARLs are generally used, one for in-control and one for out-of-control at a specified level. The ARL is a function of *h* and *k* and the underlying distribution (in-control or out-of-control) of data. The reader is encouraged to refer to Cumulative Sum Charts and Charting for Quality Improvement (Hawkins & Olwell, 1997) for a complete discussion on the methodology of best choosing these variables.

A program that operates under Windows called *Anygeth.exe* is also available for download at http://www.stat.umn.edu/cusum/software.htm. It will generate *h* and *k* values for given values of ARL, in-control means, and out of-control means. While *h* and *k* are reasonably easy to calculate for a normal distribution, they are more complex for Poisson and other distributions. The use of the software greatly simplifies the task. Minitab, a commercially available statistics program will develop both V-mask and DI CUSUM charts using default values of *h* and *k*, or the user may choose to input their own values obtained from *Anygeth.exe*. The default Minitab values are often inappropriate, and the reader is cautioned against them.

C. UNDERLYING EWMA THEORY

The use of control chart was first published in the article "Control chart tests based on geometric moving averages" (Roberts, 1959). While Shewhart charts only consider the most recent data point in testing to determine if statistical limits have been exceeded, EWMA charts consider all previous points using a weighing factor that makes the outcome more influenced by recent points. As described by Neubauer (1997):

> In brief, after multiplication by a weighting factor w, the current measurement is added to the sum of all former measurements, which is weighted with $(1 - w)$. Thus, at each time t (t = 1,2,...), the test statistic z_t [$= w\ \bar{x}_t + (1 - w)z_{t-1}$], with $w \in]0;1]$, can be obtained.[2] The computed z_t values are displayed on a control chart over the course of time. Because the mean $\bar{x}_t = (x_{1t} + x_{2t} + \ldots + x_{nt})/n$ of the n control observations per run is used, this control chart is called the EWMA- $\bar{x}$ chart. Another way of expressing this is:
>
> $$z_t = w \sum_{i=0}^{t-1} (1-w)^i \bar{x}_{t-i} + (1-w)^t z_0$$
>
> with the first value z_0 in this sum generally being set to the mean of former observations. This smoothing process means that the contribution of a value to the test statistic decays exponentially by time or by the number of new observations, with the speed of decay being adjustable by the weighting factor.
>
> The limits for warning and action of the EWMA chart differ from those of a Shewhart chart and have to be computed separately, as shown later. The EWMA control chart differs from the similar Cusum chart by using the additional weighting factor, which allows the adjustment of shift sensitivity. (Setting the EWMA weighting factor w = 1 yields a Shewhart control chart.) Because of this flexibility, the EWMA chart has drawn increasing attention in industrial quality-control practice during the past few years, as shown by the number of publications in the *Journal of Quality Technology* since 1989.

A complete description of the methodology to choose the control limits and the weighing factor is described in Chapter V, Section 4 of this thesis. The reader is encouraged to refer to "The EWMA control chart: properties and comparison with other

quality-control procedures by computer simulation" (Neubauer, 1997), for a more complete description of the theory and use of EWMA control charts.

D. APPLICATION TO AIRCRAFT MAINTENANCE AND FORECASTING

Changes in demand for aircraft repair parts can create two conditions that are not desirable. If the demand for parts increases and supply stocks are not replenished in an expeditious manner, aircraft may not be available to fly required missions due to lack of repair parts. Conversely, if demand falls for parts, scarce resources may be expended repairing more parts than are needed. If modern techniques can be shown to detect changes in demand patterns quickly enough without excessive false alarms, managers charged with keeping the right parts available to the fleet could have a new tool to more efficiently and effectively support Naval Aviation.

E. CHAPTER SUMMARY

Maintenance of naval aircraft is subject to variability as is any process. There are multiple types of variability and the tools to detect them differ. While Shewhart Xbar and R charts excel at detecting transient special cause variability, they have definite limitations. With no memory, they are not very effective in detecting small to medium size shifts in the process even if these shifts persist. CUSUM and EWMA charts excel at detecting persistent shifts even if the shifts are relatively small; in fact, among all methods with the same in-control ARL they are probably optimal to detect an out-of-control state most quickly. The DI CUSUM chart provides a concise method of implementing the CUSUM technique and can be developed quickly using commercially available software such as Minitab. Accurate determination of the *k* and *h* variables is important to ensure that the CUSUM chart sets off alarms when an unwanted shift occurs while remaining relatively free of false alarms. EWMA control charts can also be developed quickly using Minitab. Determination of the weighing factor and control limits is important in obtaining the correct sensitivity with a reasonable level of false alarms. Due to their ability to detect changes in underlying data distributions, CUSUM and EWMA can improve the ability of parts managers to detect persistent shifts in demand patterns for spare parts.

THIS PAGE INTENTIONALLY LEFT BLANK

V. RESEARCH ANALYSIS

A. INTRODUCTION

The research was performed in two phases: validation of methodologies and analysis of components. To first validate the analysis methodologies, the data set for Component Z was analyzed using CUSUM and EWMA control charts. The data set for Component Z was formed by combining two data sets each comprised of 500 data points. The first 500 data points followed a Poisson distribution with a mean of 0.5, and the second 500 data points followed a Poisson distribution with a mean of 0.7. This computer generated data set was created to follow the type of variation, a persistent shift in mean, for which a CUSUM control chart is most exactly tailored to detect. By comparing CUSUM and EWMA charts generated against a data set with known characteristics, the methodologies were validated as the resulting control charts were reflective of the characteristics of the data set and the designed ARL.

Next, CUSUM and EWMA control charts were generated for the remaining four real data sets and the results were compared to each other relative to the known variability in the data sets. Two characteristics of the control charts were evaluated:

1. How effective were each of the control chart at detecting real shifts with limited numbers of false alarms?
2. How efficient were each of the control charts as gauged by the amount of effort it required to set up and produce each chart?

B. ANALYSIS TECHNIQUES USED

1. Data Sets Analyzed

In order to compare the effectiveness of the CUSUM and EWMA methodologies to detect changes in underlying distributions, five sample data sets were chosen for analysis, namely data sets for Components “Z”, “1”, “14”, “16”, and “23.” The data set for Component “Z” was created with a known distribution and change in distribution (Poisson with a mean of 0.5 for the first 500 cases, and then with a mean of 0.7 for subsequent cases), so that the change in distribution was matched the parameters for

detection by CUSUM. Four real data sets containing the demand figures for actual naval components in operation in the fleet were also examined where the variability in demand ranged from very low to extreme. (The actual names and associated aircraft for these four real components were not included in this document in order to safeguard sensitive failure information.) These four data sets were chosen as representative of the range of variability that occurs naturally during fleet operations. By using a designed data set and four real data sets spanning the range of variability as seen in the fleet, the outcome of the analysis could be applicable to a wide range of systems in operation in the Navy.

The four real data sets were extracted from the 3M data system and were chosen to be representative of the range of variability seen within normal operations in the fleet. On one end of the spectrum, the data set from Component 1 displayed very erratic behavior with sharp "dog legs' evident in the demand for spares as seen in Figure 7.

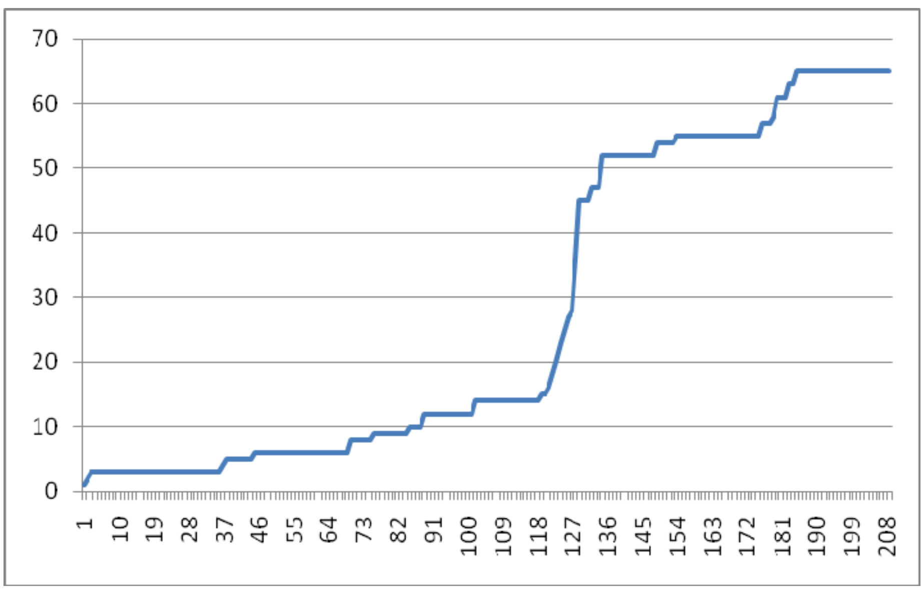

Figure 7. Cumulative BCMs for Component 1

(The term "dog leg" is often used in describing data with sharp bends as it is similar to the shape of a dog's rear leg.) On the other end of the spectrum, the data set from Component 23 displayed relatively stable behavior with minor levels of variability in the demand for spares. In between, two data sets from Components 14 and 16 were

chosen with intermediate levels of variability. It should be noted that none of the real data sets analyzed displayed the "classic" persistent shift in mean for which CUSUM is optimally designed.

2. Method of Analysis

For the created data set for Component Z, CUSUM and EWMA charts were generated using Minitab. Refinements were made to the methodology until the outputs of the charts were reflective of the data set itself with false alarms in the range predicted by the in-control ARL settings and frequent alarms soon after the mean shift at point 501.

For each of the four real datasets that were chosen using the methodology described above, three charts were developed:

The first chart displays the cumulative BCMs by week for the component over the period of analysis and was generated using Excel 2007. This provided a visual depiction of the demand variability over time.

The second chart displays a CUSUM chart as generated by Minitab for each component using *h* and *k* values as determined using *Anygeth.exe.* For the CUSUM charts, the in-control ARL was set to 100 weeks. As such, one would expect to see about one false alarm every 100 weeks.

The third chart displays a EWMA control chart for the data set up to have an ARL of 100 weeks. Similarly, one would expect to see one false alarm every 100 weeks.

By examining the CUSUM and the EWMA chart, and comparing them to the shifts in the cumulative BCM chart, the author will assess the efficiency and effectiveness of each method relative to its ability to determining shifts in demand without excessive false alarms being present.

3. Validation of CUSUM Methodology and Modification of Variables between *Anygeth.exe* and Minitab

When attempting the CUSUM validation phase of this effort against the Component Z sample data set using *h* and *k* values obtained from *Anygeth.exe* within the Minitab program, several problems arose. This section will discuss the steps necessary to successfully integrate the two programs.

Before one can determine the parameters of the CUSUM scheme, one must first select an in-control ARL. Recall that this specifies the average number of time periods until one gets a false alarm while in control. Longer ARLs imply fewer false alarms, but can result in a longer delay until a true alarm is signaled. The second parameter needed is to select the out-of-control mean. This is set based on the context of the problem, and is generally based upon the size of a change that starts to have serious practical effects. Larger departures are more robust against model misspecification (i.e., the distribution is not exactly as assumed).

Using *anygeth.exe*, one can find the values of h and k that are tuned for the size of the departure desired. A screen shot of the dialog for the example with the Z Component showing the determination of the parameters for the CUSUM for a Poisson distribution with mean shift from 0.5 to 0.7 and an ARL of 100 is displayed in Figure 8.

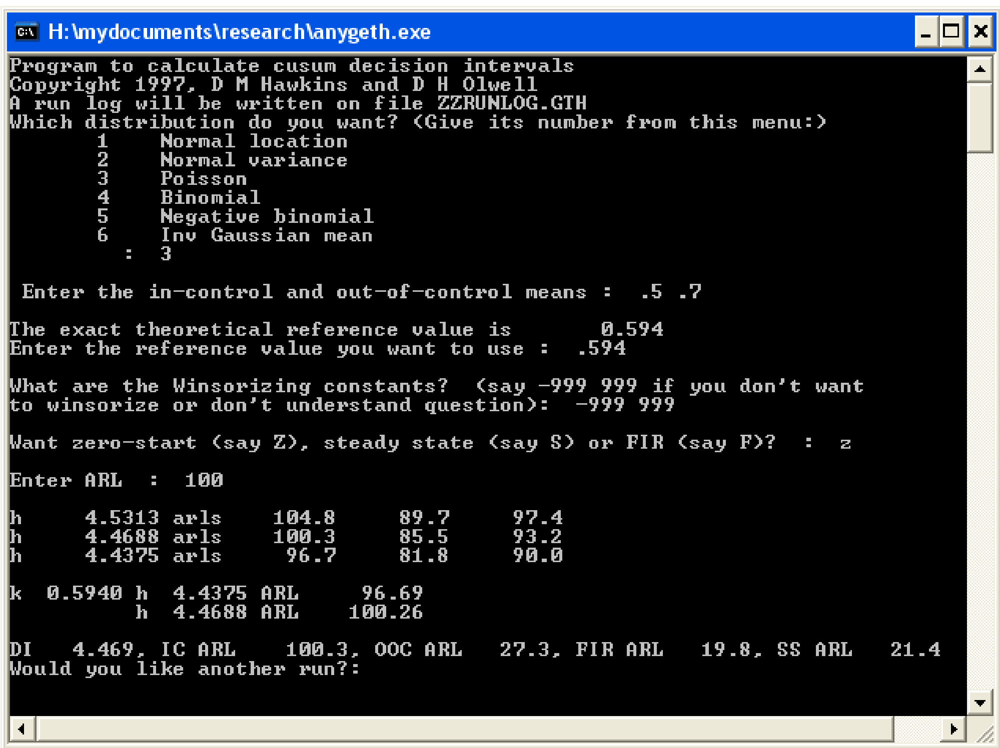

Figure 8. Dialog from *Anygeth.exe*

Minitab and *anygeth.exe* use slightly different parameterizations, so one must convert between them. The following instructions should be followed carefully.

1. Calculate h and k using Anygeth.exe for the desired ARL and out of control state. See Figure 8 for an example. From Figure 8 we obtain $h = 4.4688$ and $k = 0.5940$.

2. Determine the value of $\hat{k} = k -$ *in-control-mean-* and insert into the CUSUM Plan k field. Do not use the unadjusted value of k from anygeth.exe. For our example, we get $\hat{k} = 0.5940 - 0.5 = .0940$. See Figure 9.

3. Insert the value of h into the CUSUM Plan h field. Here you do use the unadjusted value. See Figure 9.

4. Change the value of the Standard Deviation to 1.0. See Figure 10.

5. Insert the value of the in-control mean into the CUSUM Target field. See Figure 11.

6. After entering all information above, run the CUSUM by clicking "OK". See Figure 11.

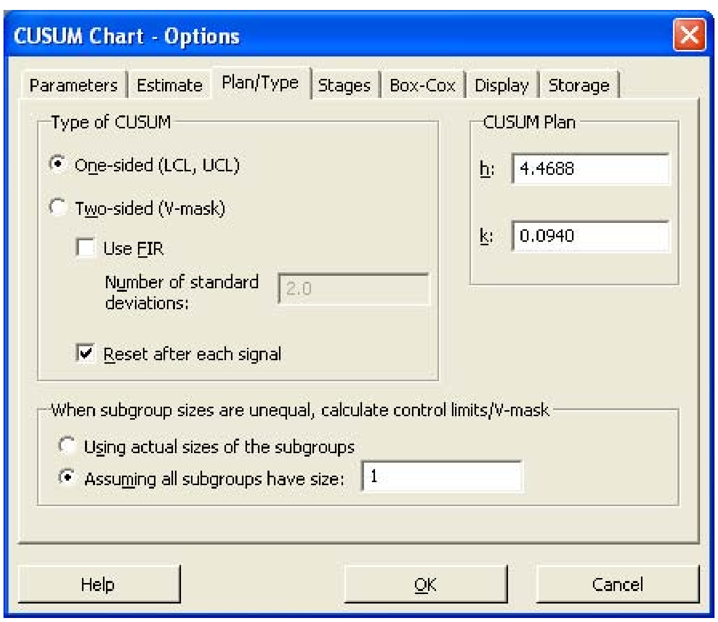

Figure 9. CUSUM Chart – Options Dialog Box – Plan/Type Tab

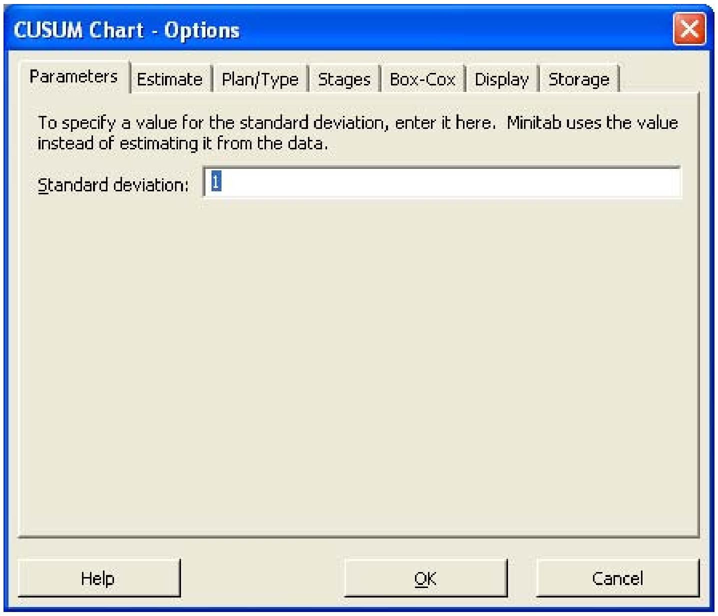

Figure 10. CUSUM Chart – Options – Parameters Tab

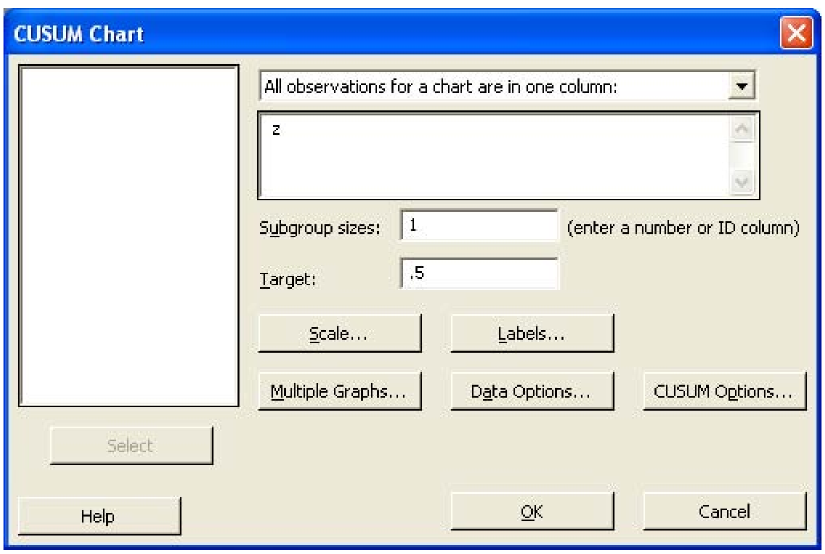

Figure 11. CUSUM Chart Dialog Box

The CUSUM chart generated using the above values of *h* and *k* as determined by Anygeth.exe, 4.4688 and 0.594, respectively, is shown in Figure 12. The chart produced alarms at: 29, 53, 167, 185, 221, 284, 329, 527, 553, 566, 584, 637, 691, 755, 771, 797, 828, 843, 867, 908, 931, 974, and 996. This equates to seven false alarms during the in-control period which is well within normal variation for an ARL of 100 with 500 data

points in the in-control region in which five false alarms would be the average number expected. The CUSUM chart accurately detected the real shift in mean first at point 527 and provided frequent alarms afterward providing validation of the methodology employed to set up Minitab using *Anygeth.exe*. Note that Minitab was set to reset after each alarm. The output of this chart against the known data set for Component Z provided a validation of the methodology used to create the chart using *Anygeth.exe* and Minitab.

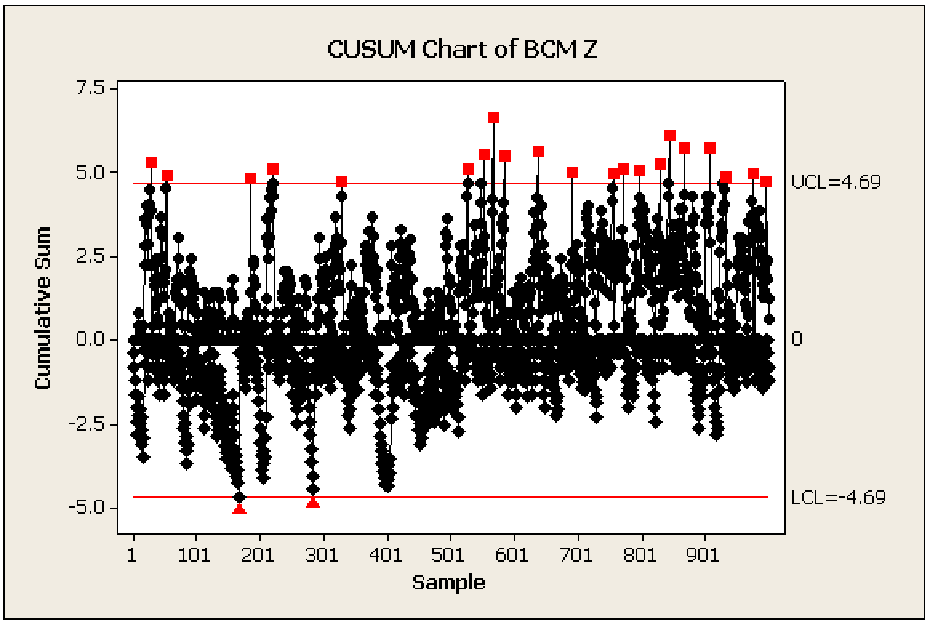

Figure 12. CUSUM Chart of Component Z with ARL set to 100

4. Validation of Methodology for EWMA

In order to accurately compare EWMA and CUSUM techniques, they must be set up on a "level playing field." Both must be designed with the same ARL, or the results cannot be compared fairly. In this section, the methodology to set up the Minitab EWMA control chart function to have an ARL of 100 will be described and validated against the Component Z data set.

Neubauer (1997) in the article "The EWMA control chart: properties and comparison with other quality-control procedures by computer simulation," described a graphical method for determining the ARL for an EWMA control chart. As seen in

Figure 13, one can obtain the weighting factor "*w*" as a function of the desired shift sensitivity of the chart. For a shift of one standard deviation, a weighing factor of 0.15 can be obtained from the chart.

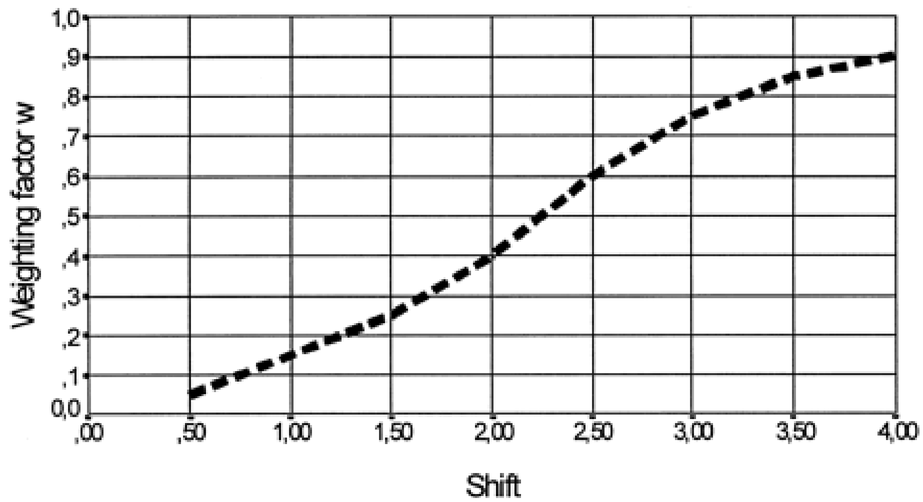

Figure 13. Optimal *w* for EWMA charts according to the shift *d* from Neubauer 1997

For a desired ARL of 100 and a weighting factor of 0.15, one can obtain the limit *q* value of 2.3 as seen in Figure 14.

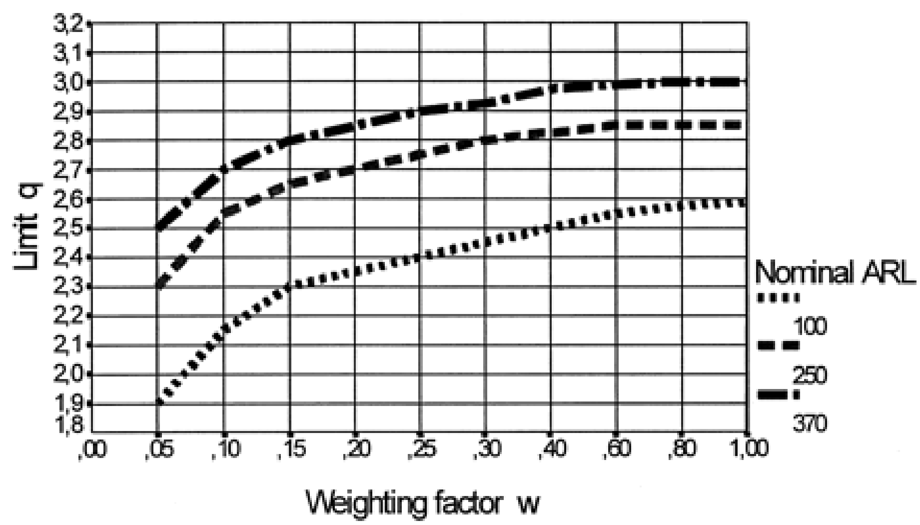

Figure 14. Determining the limit q of the EWMA chart after selection of w and the nominal ARL from Neubauer 1997

In order to set up Minitab to run an EWMA control chart with an ARL of 100, all that is required is to: enter the value of w = 0.15 obtained from Figure 13 in the “Weight of EWMA” field as seen in Figure 15; and to enter the limit q value of 2.3 obtained in Figure 14 in the “Display control limits at” dialogue box as seen in Figure 16. Additionally, if the in-control mean is known, as it is for Component Z, then more accurate results are obtained if that value is entered into the “Mean” field on the EWMA Chart – Options dialogue box under the “Parameters” tab. Otherwise, Minitab will calculate a mean for the entire data set.

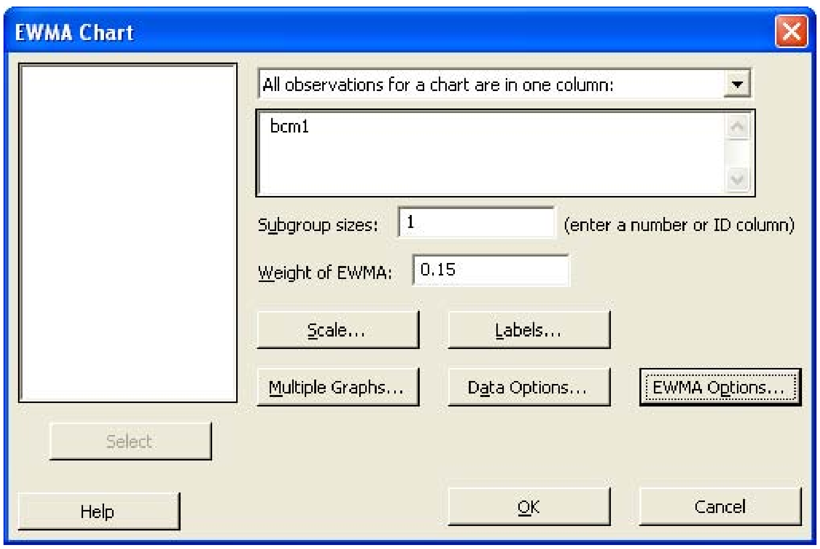

Figure 15. Minitab EWMA Chart Dialogue Box

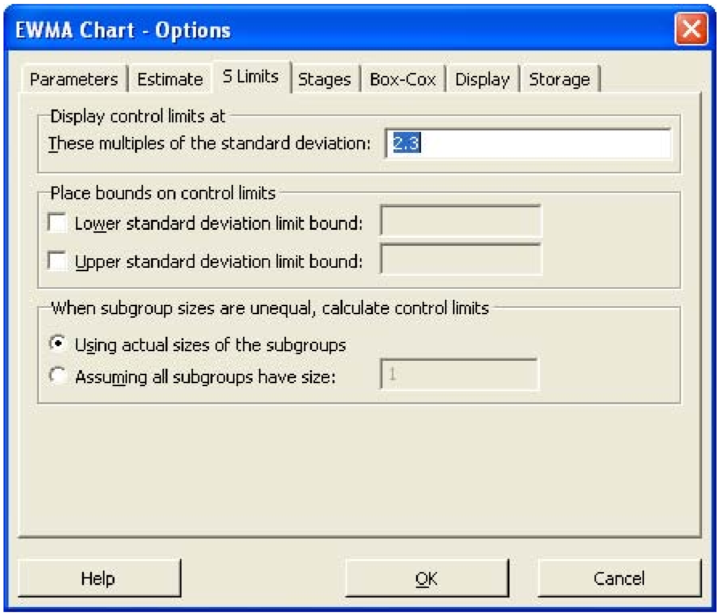

Figure 16. Minitab EWMA Chart - Options Dialogue Box

Unlike CUSUM charts, where variables have to be recalculated for every chart requiring a great deal of user knowledge and assistance from a computer program, the EWMA chart has the distinct advantage that the set up is a one-time event using easily understood graphical techniques. For a given in-control ARL, the EWMA has the same parameters for all datasets regardless of distribution due in part to the central limit theorem.

As a validation of this methodology for determining the variables to set up Minitab to run EWMA control charts, the data set for Component Z was run in Minitab. Figure 17 displays the output.

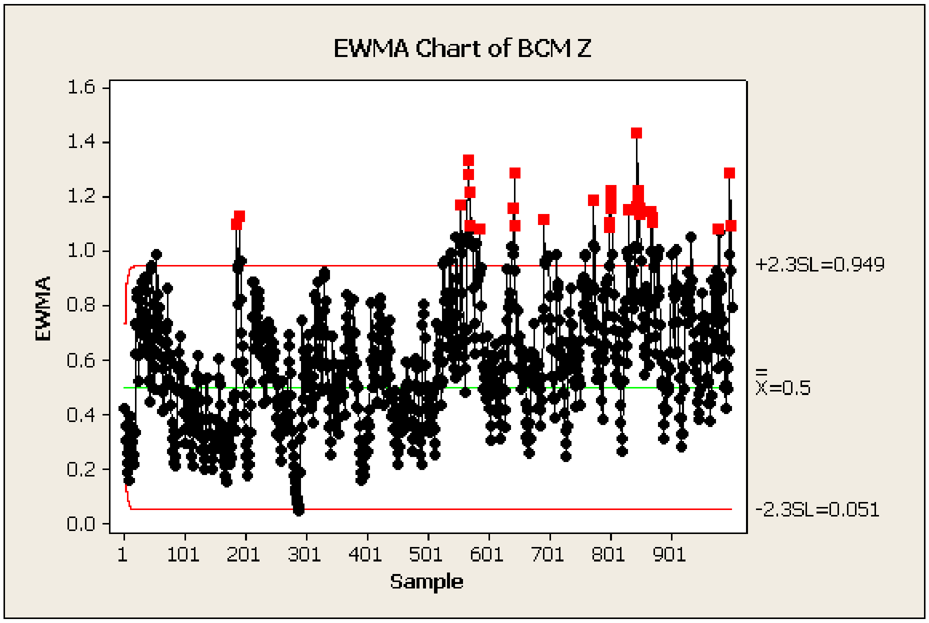

Figure 17. EWMA Chart for Component Z

Minitab displayed alarms at points 185, 189, 554, 566, 567, 568, 570, 584, 640, 642, 643, 691, 773, 798, 799, 800, 801, 802, 830, 843, 844, 845, 846, 848, 849, 868, 869, 870, 978, 997, and 998. In this case, there were two false alarms during the in-control period with an expected average of five false alarms with an ARL of 100 across 500 in-control data points. This is also within normal variation. The EWMA chart provided the first valid alarm at point 554 (27 points later than the CUSUM chart) and then alarmed consistently after that during the out-of-control region. This is exactly the behavior that would be expected from a valid EWMA chart on the data set from Component Z. Thus the process described in this section produces a valid EWMA chart when run against a data set with known characteristics.

C. COMPONENT ANALYSIS

With the methodology validated to create CUSUM and EWMA charts using Minitab and Anygeth.exe, charts were prepared for each of the four real components and the results compared and analyzed as follows.

1. Results of Analysis for Component 1

Figure 18 displays the cumulative BCMs for Component 1 on the *y* axis plotted against weeks on the *x* axis.

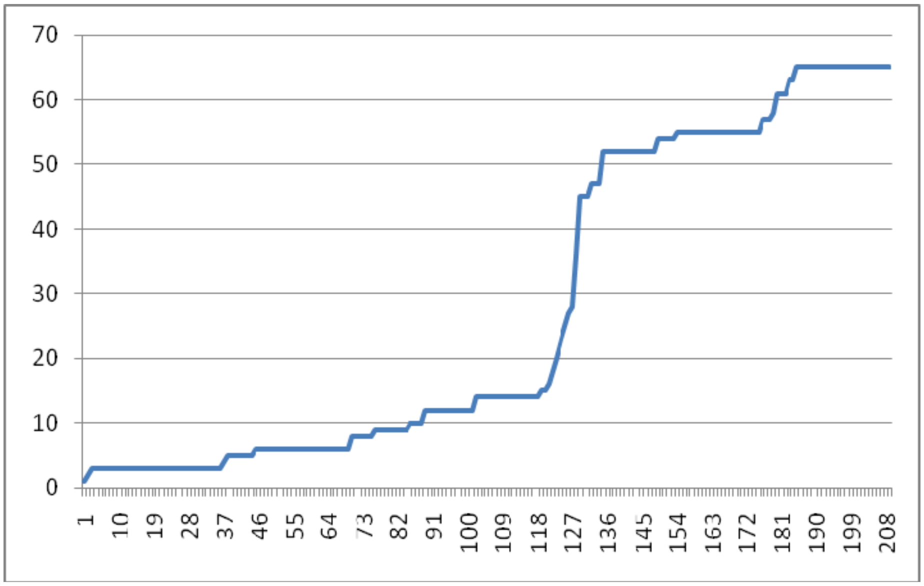

Figure 18. Cumulative BCMs for Component 1

The reader will note a relatively stable demand pattern until approximately time period 120 when the demand for Component 1 was subject to a very dramatic increase. After that, the demand returned to a relatively stable pattern until around week 180 when it spiked again only to level out again around week 190. These sharp bends in demand patterns are often described as "dog legs" as discussed earlier. Even during the relatively stable periods, there were extended plateaus or periods of zero demand.

The underlying data (BCMs/week) has a mean of 0.311 with a standard deviation of 1.049. A Poisson distribution for a data set with this mean should have a standard deviation of 0.558 indicating that this entire distribution does not follow a Poisson distribution closely. However for the period during weeks 1-119, the mean is 0.12 and the standard deviation is 0.40. A Poisson distribution would have an ideal standard deviation of 0.36 that is reasonably close to the real value of 0.40 indicating that this data follows a Poisson distribution relatively closely during that time frame and that the Poisson model can be used for analysis.

Figure 19 displays the CUSUM chart for Component 1 using a value of the mean of 0.12, which represents the average number of BCMs for weeks 1–119, a period in which the demand was reasonably stable. The in-control ARL was set at 103. Anygeth.exe provided values of k and h of 0.16 and 2.4 respectively. After adjustment, $\hat{k}$ was set at 0.04.

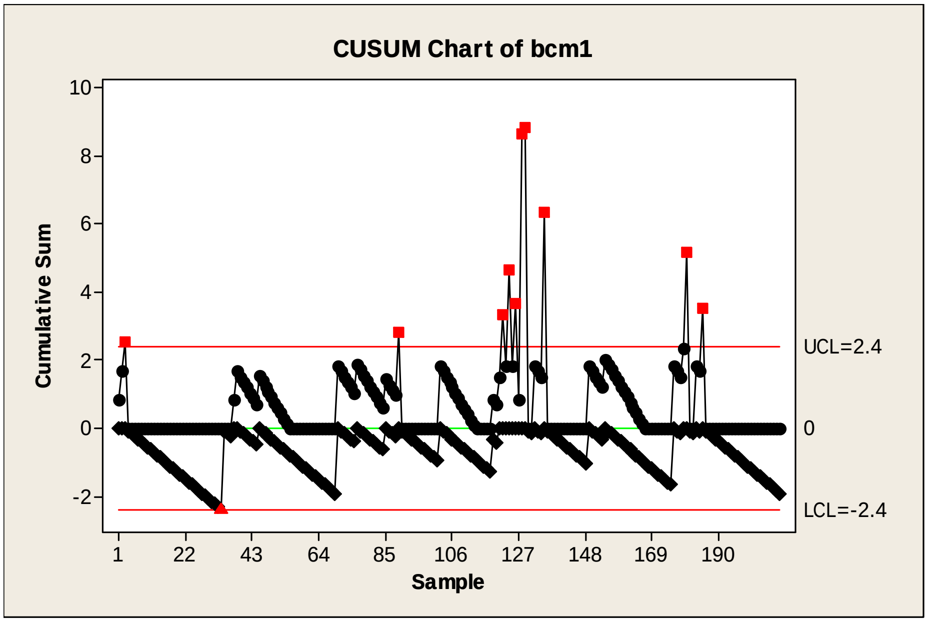

Figure 19. CUSUM Chart for Component 1

There were alarms at points 3, 33, 89, 122, 124, 126, 128, 129, 135, 180, and 185. The alarms coincide with the appearance of "spikes" and "plateaus" in the demand as seen in Figure 18.

Figure 20 displays the EWMA chart for Component 1 with an ARL of 100, weight of 0.15, and control limits at 2.3 standard deviations.

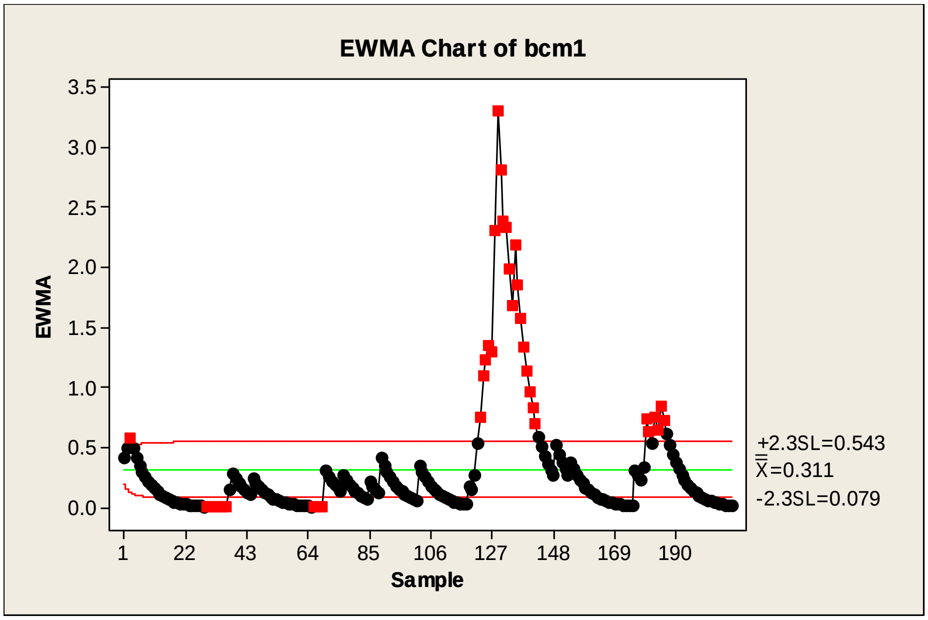

Figure 20. EWMA Chart for Component 1

The EWMA chart provided alarms at points 3, 29, 30, 31, 32, 33, 34, 35, 36, 66, 67, 68, 69, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 180, 181, 183, 184, 185, and 186. There are corresponding spikes and plateaus in the demand patterns that coincide with these alarms.

Observations: For components with relatively stable demand patterns with intermittent large spikes, CUSUM and EWMA are effective in detecting the spikes without excessive false alarms. The EWMA chart detected a movement down in demand at week 66 that was missed by the CUSUM chart. However, the reader will notice that the CUSUM chart was near alarm at that time. Overall, both chart types appeared effective in detecting shifts in demand in Component 1. The EWMA methodology requires less overhead and expertise than CUSUM for these types of components and may be a more efficient choice for monitoring these types of components.

2. Results of Analysis for Component 16

Figure 21 displays the cumulative BCMs for Component 16 on the y axis plotted against weeks on the x axis.

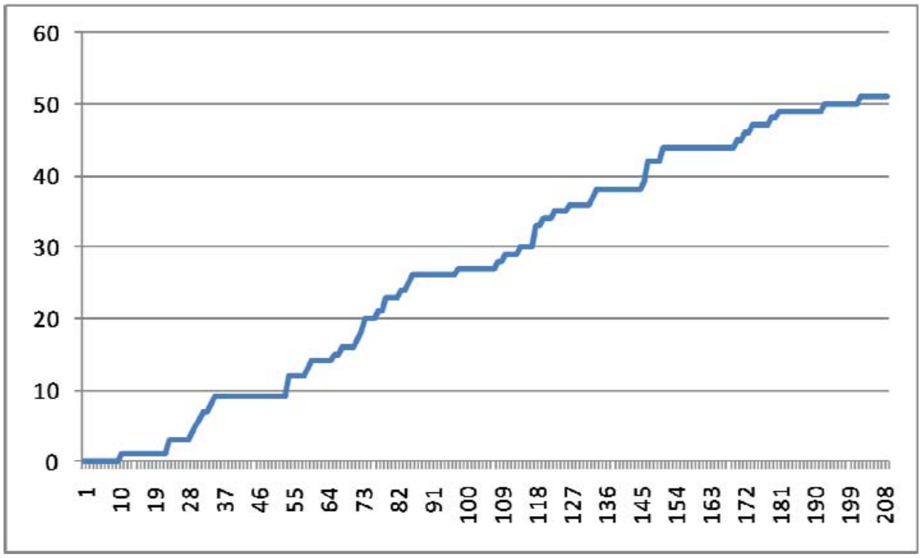

Figure 21. Cumulative BCMs for Component 16

The reader will note that the pattern is more stable than Component 1, but that it still contains spikes and plateaus in the demand pattern. The underlying data (BCMs/week) has a mean of 0.244 with a standard deviation of 0.557. A Poisson distribution for a data set with this mean should have a standard deviation of 0.494 indicating that this distribution is close, but does not exactly follow a Poisson distribution.

Figure 22 displays the CUSUM chart for Component 16 using values of *h* and *k* as determined by *Anygeth.exe* with an in-control mean of 0.244, an out-of-control mean of 0.366, $\hat{k}$ = 0.03, *h* = 3.4 and in-control ARL = 103.

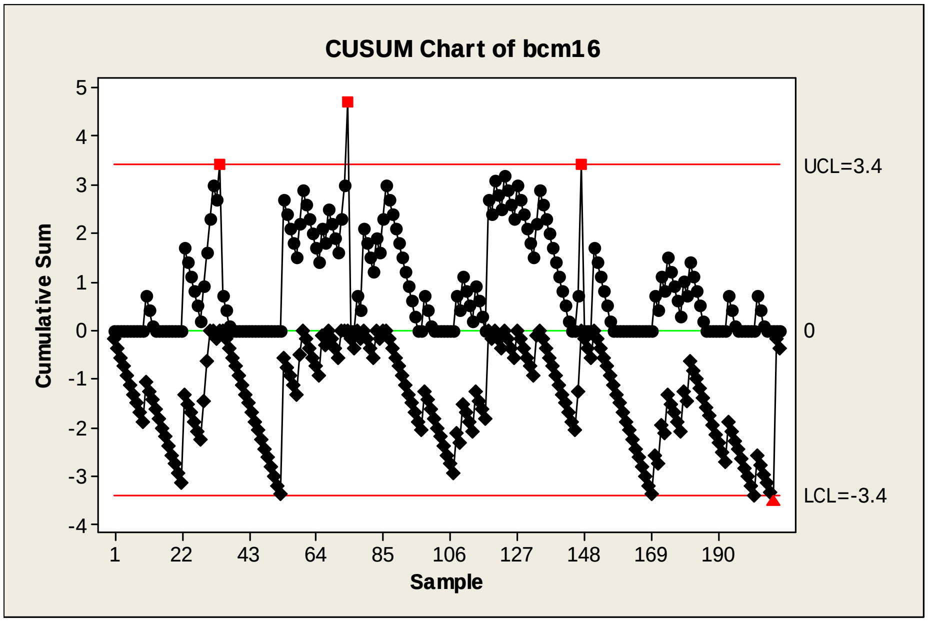

Figure 22. CUSUM Chart for Component 16

There were four alarm signals at points 34, 74, 147, and 209. With an ARL of 103, these may just be usual variation.

Figure 23 displays the EWMA chart for Component 16 with an ARL of 100, weight of 0.15, and control limits at 2.3 standard deviations.

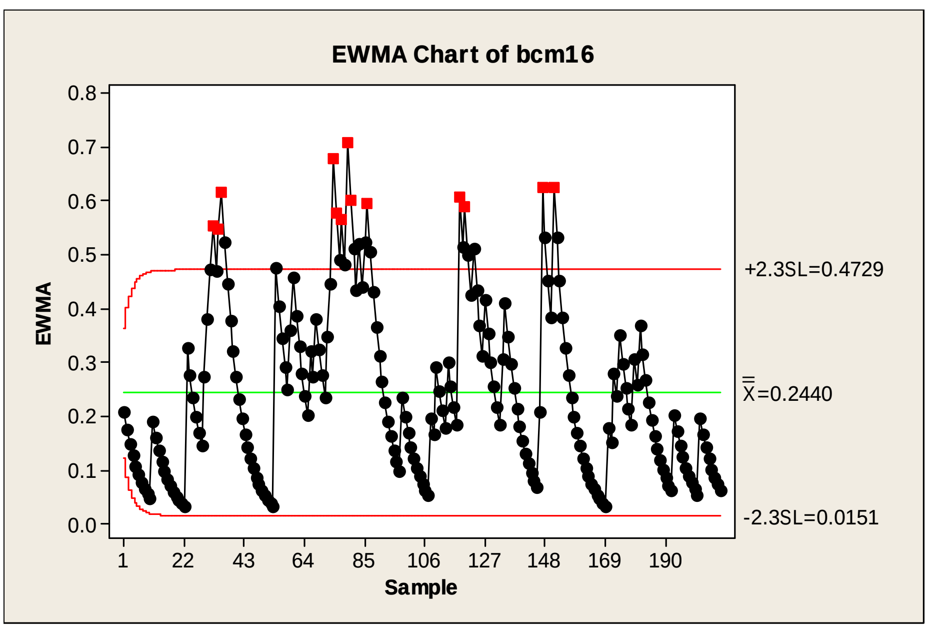

Figure 23. EWMA Chart for Component 16

In this chart, there are four groups of alarm signals, with three roughly at the same places as three of the signals on the CUSUM chart. The signals were at points 32, 34, 35, 74, 75, 77, 79, 80, 86, 118, 120, 147 and 151. The EWMA chart signaled an alarm around data point 120, where the CUSUM did not quite signal, while the CUSUM chart signaled an alarm at the very end of the data.

Observations: Component 16 has a demand pattern that consists of several intermediate level spikes followed by plateaus. The alarms of each tool did correspond to periods of spikes and plateaus in the data. The EWMA methodology requires less set up time than CUSUM for these types of components and may be a more efficient choice for monitoring these types of components.

3. Results of Analysis for Component 14

Figure 24 displays the cumulative BCMs on the *y* axis versus weeks on the *x* axis for Component 14.

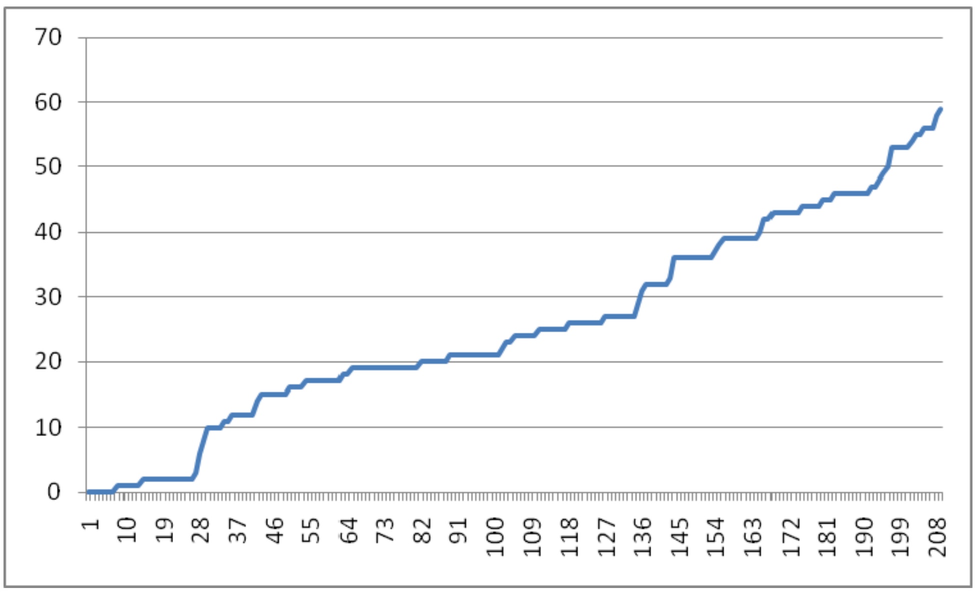

Figure 24. Cumulative BCMs for Component 14

The reader may note that the pattern is similar to Component 16, and contains spikes and plateaus in the demand pattern. This data sample has a mean of 0.282 with a standard deviation of 0.592. A Poisson distribution for a data set with this mean should

have a standard deviation of 0.531 indicating that this distribution indicating that this distribution while closer to a Poisson than Component 1, does not exactly follow a Poisson distribution.

Figure 25 displays the CUSUM chart for Component 14 using a $\hat{k}$ hat value of 0.0657 and an *h* value of 3.53 as determined using Anygeth.exe with an ARL of 100.7. The in-control mean was 0.282 and the out-of-control mean was 0.423. There were alarms at points 26, 29, 78, 131, 137 and 197.

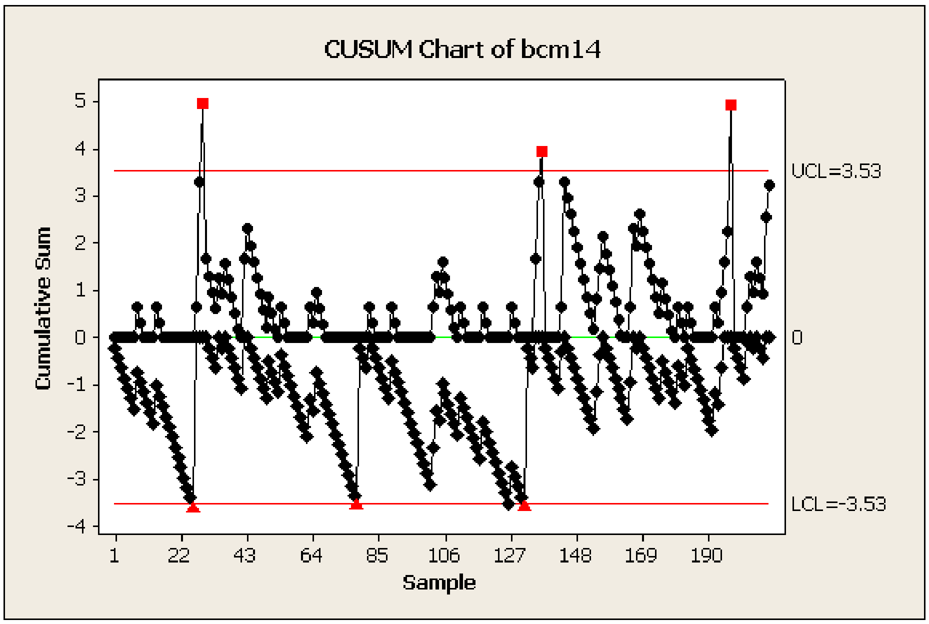

Figure 25. CUSUM Chart for Component 14

Figure 26 displays the EWMA for Component 14 with an ARL of 100, weight of 0.15, and control limits at 2.3 standard deviations.

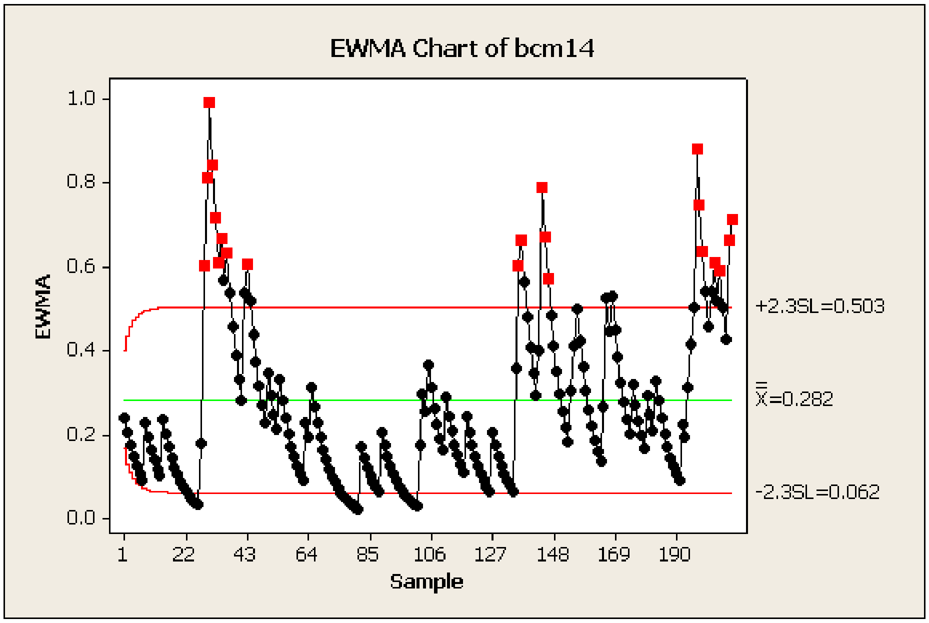

Figure 26. EWMA Chart for Component 14

There were alarms in three groups at points 28, 29, 30, 31, 32, 33, 34, 36, 43, 136, 137, 144, 145,146, 197, 198, 199, 203, 205, 208 and 209. The positive alarms were nearly the same location as the CUSUM alarms but slightly slower in two of three cases. The EWMA chart did not produce alarms in the negative direction even though it came very close three times at the same places that the CUSUM chart alarmed.

Observations: Component 14, much like Component 16 has a demand pattern that consists of several intermediate level spikes followed by plateaus. Each tool triggered at the location of spikes or plateaus in the demand patterns. CUSUM was slightly faster and in this case appeared slightly more sensitive to the negative alarms after plateau periods. The EWMA methodology requires less overhead than CUSUM for these types of components and may be a more efficient choice for monitoring these types of components.

4. Results of Analysis for Component 23

Figure 27 displays the cumulative BCMs plotted on the y axis versus weeks on the x axis for Component 23.

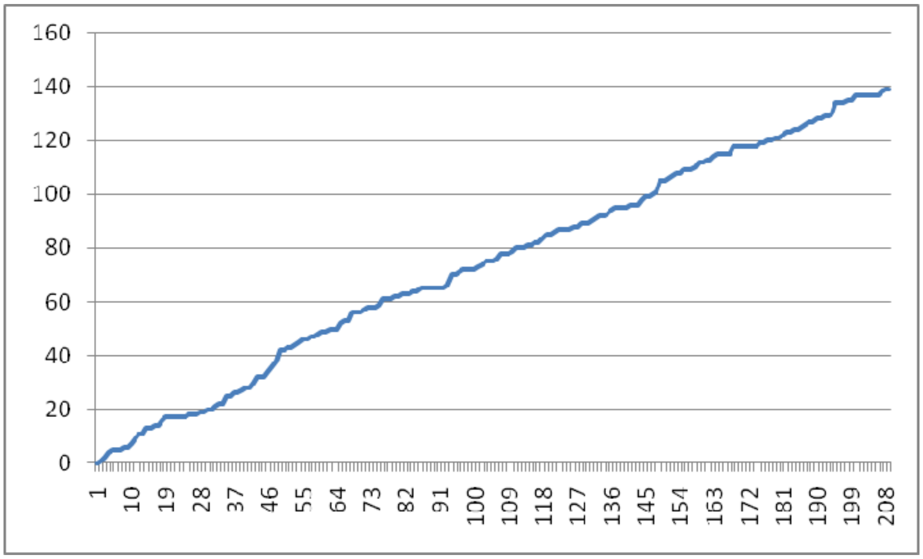

Figure 27. Cumulative BCMs for Component 23

The reader will note that the demand pattern is more stable than the other components. This component was chosen to represent the stable end of the variability spectrum. This data sample has a mean of 0.65 with a standard deviation of 0.84. A Poisson distribution for a data set with this mean should have a standard deviation of 0.81 indicating that this distribution is closer to a Poisson distribution than the other components discussed.

Figure 28 displays the CUSUM chart for Component 23 using a $\hat{k}$ hat value of 0.16, an h value of 4.43 as determined using Anygeth.exe with an ARL of 100. The in-control mean was 0.65 and the out-of-control mean was 1.0.

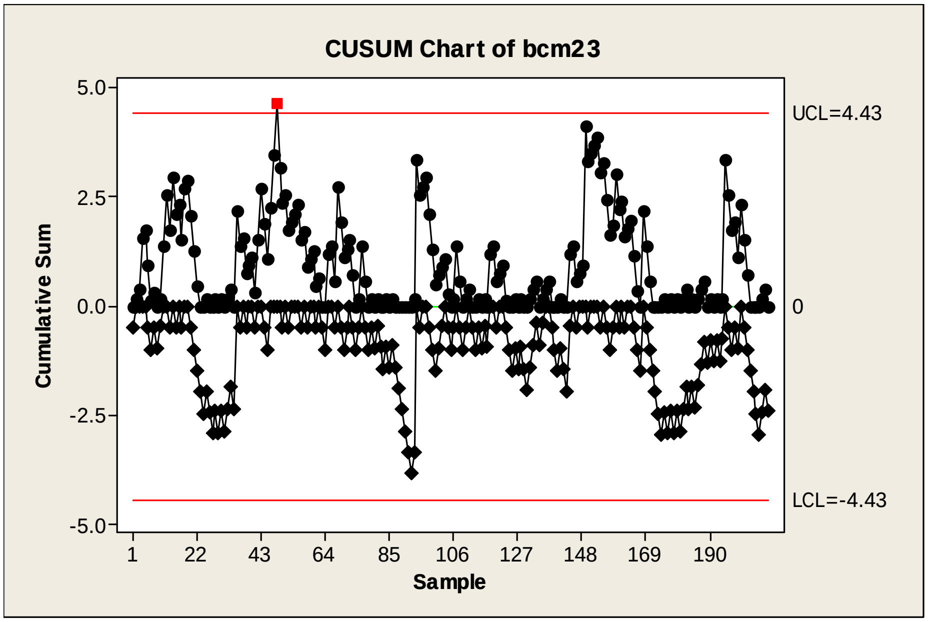

Figure 28. CUSUM Chart for Component 23

There is a single alarm signal at point 48. With an ARL of 100 across 209 data points, this alarm is mostly likely just due to normal variation.

Figure 29 displays the EWMA for Component 23 with an ARL of 100, weight of 0.15, and control limits at 2.3 standard deviations.

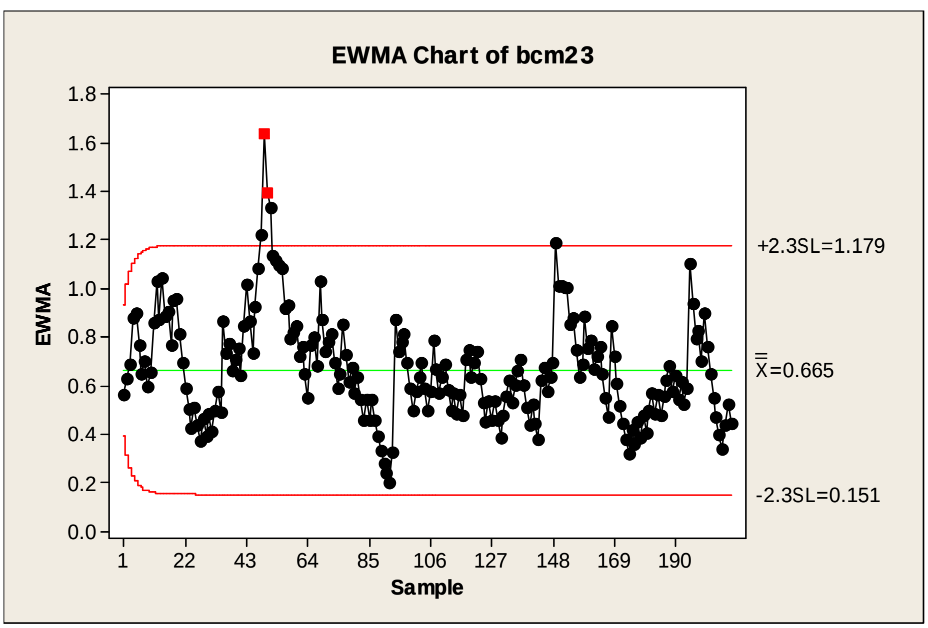

Figure 29. EWMA Chart for Component 23

There are two alarms at points 49 and 50. With an ARL of 100, this again is probably just normal variation.

Observations: Both methods picked up the small spike in demand around week 49 with CUSUM alarming one week earlier. Of all of the components studied, this one had the least activity in alarms, which is consistent with an in-control process with very little variation. While both the CUSUM and EWMA methods were effective for this type of component, the EWMA method was more efficient.

D. GENERALIZATION OF FINDINGS AND RECOMMENDATIONS

Both the CUSUM and the EWMA methodologies were very capable of detecting spikes and plateaus in the demand patterns when correctly set up to analyze the data. In general, CUSUM provides alarms slightly faster than EWMA. However, the CUSUM technique required a unique set up for each component requiring user expertise and the use of a computer program to generate required variables. EWMA has the advantage of a single set up for all components with the same ARL, using an easily understood graphical approach. Due to the central limit theorem, the EWMA method is relatively insensitive to the underlying distribution of the data while the CUSUM methodology does require knowledge of the underlying distribution. For most applications monitoring aircraft component spares usage, EWMA appears to be the most efficient method with very little loss in efficiency.

E. CHAPTER SUMMARY

In this chapter, the methodology to detect shifts in underlying demand distribution for aircraft components was first validated on a data set for an imaginary Component Z. The processes to set up the required variables needed to use the Minitab program to produce CUSUM and EWMA charts were developed and validated.

Control charts were developed using CUSUM and EWMA for four components spanning a wide range of variability. The effectiveness of the charts relative to determining shifts in underlying trends and the efficiency of the charts relative to the time and expertise required to generate the charts were compared.

Both methods were very capable at detecting changes in the demand patterns when properly configured within Minitab. CUSUM appeared to be slightly faster in detecting changes, while EWMA was much easier to set up across multiple data sets. While CUSUM requires a custom set up for each different component with an understanding of the underlying data distribution, EWMA requires a single graphical set up and is insensitive to the underlying data distribution. For most component applications within the Navy, EWMA is a more efficient tool with only a slight loss of effectiveness.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. APPLICATION OF STUDY

A. INTRODUCTION

The use of CUSUM and EWMA techniques, as applied to detection of demand patterns for WRAs within the Naval Air Forces, has been examined. When properly configured, both of these methods have been shown to be effective in determining changes in underlying demand patterns with the EWMA being the more efficient of the two. While this thesis was focused on components from aircraft, the methods developed in this paper would have broad general application to any process where the user desired to know if the underlying data distribution has changed. In this chapter, the author will discuss what steps will be required for the reader to apply this study to components or processes as desired.

B. RECOMMENDATIONS – STEPS NEEDED TO APPLY

As with any management effort, the reader who wishes to apply these techniques should determine the cost and benefit to do so. By first focusing on those items that are high cost or high risk first, the most efficient use of resources can be assured. After priority ranking the items or processes to be monitored, the following steps are suggested:

1. Determine what data is representative of the process that is to be monitored and if that data can be obtained with a reasonable level of effort. A sufficient amount of historical data must be gathered to have an understanding of an in-control mean and how much variation there is in the data before these tools can be effectively used. A minimum of 30 data points is recommended.

2. Obtain access to Minitab. In many areas of the Department of Defense, this program is already licensed for use. The reader may also purchase Minitab. (Other programs may be as effective in developing these charts, but the steps in developing and inputting the control variables would require revalidation. There are several ways to parameterize a CUSUM and EWMA, and one must attend carefully to the notation used by a given piece of software.)

3. Determine if EWMA or CUSUM charts are to be developed. EWMA charts require less time and expertise to use, but CUSUM charts can be faster to alarm especially where the size of the change is known beforehand.

4. For CUSUM charts, the user will require *Anygeth.exe*. This Windows based program is available for free download at http://www.stat.umn.edu/cusum/software.htm.

5. Follow the steps in Chapter V to set up Minitab with the correct variables and develop the control charts.

6. Monitor the process and adjust the variables as needed. Longer ARL lengths will give fewer false alarms at the price of slower reaction.

7. Consider the overall system false alarm rate when setting component ARLs. For example, to have an ARL of 100 for a system managing four components, one would not use an ARL of 100 for each component, as that would result in an ARL of 25 for the set. Rather, one could simply use an ARL of 400 for each of the four components, leading to a system ARL of 100. Other allocations of ARL besides equal distribution are possible among components to get a desired system ARL. (Serel, Moskowitz, Tang, 2000)

C. CHAPTER SUMMARY

EWMA and CUSUM techniques can provide alarms when the underlying data distribution changes. This capability could be very valuable in a wide range of applications. Effort is best applied to high cost or high risk items where representative data can be obtained. By following the steps in this section, application of the findings presented in this thesis could be far ranging.

VII. CONCLUSIONS

A. KEY POINTS AND RECOMMENDATIONS

Maintaining aircraft in a high state of readiness around the globe is a large challenge for the Navy. In order to keep the aircraft repaired and ready for missions, managers must allocate resources to keep spare parts available for use by maintainers when and where needed. Predictions for future demand are made using historical demand patterns. If the underlying data distributions of the spare requirements change, the predictions can be in error, leading to oversupply conditions if demand is falling or lack of spare parts needed to keep aircraft operational if demand is increasing. Thus it is very important for managers to know if the underlying demand distribution of the spares requirements is changing.

In this thesis, the author:

- Examined the best data to monitor to track demand;
- Validated the use of CUSUM and EWMA on a designed data set;
- Developed a process to generate CUSUM and EWMA control charts using widely available software tools;
- Analyzed data using CUSUM and EWMA from four components with a range of variability spanning the range typically found in fleet operational equipment; and
- Made recommendations about the relative efficiency and effectiveness of the CUSUM and EWMA techniques.

BCM data was found to be very representative of wholesale demand and can be easily obtained from the Navy 3M system. CUSUM and EWMA were both validated against a designed data set and processes were developed to determine the correct values of variables required to set up these tools in Minitab. Both CUSUM and EWMA were capable of detecting shifts in demand data. CUSUM generally provide a faster alarm, but required considerably more expertise and time to use. A unique set up along with an understanding the underlying data distribution was required for each component when developing CUSUM charts. EWMA charts tended to be slightly slower in providing

alarms, but were much easier to set up with a single set up required for all components for a given ARL. Overall, EWMA charts are more efficient to use with a slight loss in effectiveness.

B. AREAS TO CONDUCT FURTHER RESEARCH

While the use of CUSUM and EWMA, as described in this thesis, demonstrate the ability to detect changes in underlying data distributions using widely available tools, there is much more that could be done to expand the potential benefits.

1. The process as described still requires a considerable amount of human intervention. BCM data must be extracted and loaded into a Minitab program manually. Each component would require individual monitoring. If an integrated system was designed that would automatically load the BCM data into a computer system with EWMA tools running the background, alarms could be set up for a given ARL simply. Early warning could thus be provided to inventory, engineering, logistics, and program managers for a very large number of components with no human intervention. These alarms could allow managers to make mitigating actions before aircraft readiness was impacted.

2. While the above system would provide alarm and detection of changes in underlying demand, it would not analyze the cause of the change. In discussions with senior managers at NAVICP, they described the "holy grail" in demand for spares analysis. This is the ability to mine data using variables such as BCM rates, removals, Turn-Around-Times, and then diagnose the nature and causes of changes in demand patterns. Thus a manager would not only know that something is changing, but would also have insight into the cause of the change. For example, if Removal, Repair and/or BCM rates are changing at a site, but not at other sites doing similar work, this indicates a potential support issue at the site where the change is occurring. If site data for sites hosting similar aircraft could be combined, the combined demand data could provide early warning of an emerging change in failure mode on an

item before the change is apparent at any one site. Historically these approaches have floundered due to inability to get real time data and lack of proper heuristics driving too many false alarms. The integration of multiple data sets and the knowledge mining from these data sets would be a very valuable place for continued research.

3. Care should be taken to set ARLs for components, regardless of whether a CUSUM or EWMA chart is used, to avoid a systemic rate of false alarms that is intolerable to management. Further research into the allocation of ARL among components based on economic factors unique to NAVAIR and the Naval Aviation Enterprise would be useful.

C. SUMMARY

The Navy is charged with maintaining combat ready aircraft all around the globe prepared to go into harm's way to defend the interests of our country. These missions can only be completed with aircraft that are properly maintained. Spare parts must be available when and where needed and this requires projection of future demand. If the demand patterns are changing, for any reason, managers need to know about the change in time to allow mitigating actions such that the readiness of our aircraft is sustained.

In this thesis, the author has demonstrated that the CUSUM and EWMA methods are very capable of detecting changes in underlying distribution patterns. The processes to prepare CUSUM and EWMA charts using widely available software tools were developed and validated. The CUSUM charts are very effective at providing early alarm of changes but do require specialized knowledge and additional software tools to effectively use. EWMA charts are nearly as effective and require much less time and skill to use. With the use of these tools, Navy managers could take a more proactive response to issues enabling more aircraft to be in a state of combat readiness.

While this thesis focused on a specific issue of spares for Navy aircraft, the concepts and methodologies developed within this thesis would readily apply to any process for which the user wanted to detect changes in sufficient time to allow mitigating actions. Recommendations were made for further research that could automate these processes and provide more information about how to solve the specific issue in addition to detecting the issue.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bower, K. M. (n.d.). *Using CUSUM Charts to Detect Small Process Shifts.* Retrieved July 6, 2010, from Minitab.com: http://www.minitab.com/uploadedFiles/Shared_Resources/Documents/Articles/cumulative_sum_charts_for_small_shifts.pdf

Croston, J. D. (1972). Forecasting and Stock Control for Intermittent Demands. *Operational Research Quarterly Vol 23, No 3* , 289–303.

Eaves, A. H., & Kingsman, B. G. (2004). Forecasting for the Ordering and Stock-Holding of Spare Parts. *The Journal of the Operational Research Society, Vol 55, No 4* , 431–437.

Gardner, E. S. (1985). Exponential Smoothing: The State of the Art. *Journal of Forecasting Vol 4* , 1–28.

Ghobbar, A. A., & Friend, C. H. (2003). Evaluation of forecasting methods for intermittent parts demand in the field of aviation, a predictive model. *Computers and Operations Research* , 2097–2114.

Hawkins, D. M., & Olwell, D. H. (1997). *Cumulative Sum Charts and Charting for Quality Improvement.* New York Berlin Heidelberg: Springer-Verlag.

Neubauer, A. S. (1997). The EWMA control chart: properties and comparison with other quality-control procedures by computer simulation. *Clinical Chemistry , 43*, 594–601.

Roberts, S. V. (1959). Control chart tests based on geometric moving averages. *Technometrics* (1), 239–250.

Serel, D. A., Moskowitz, H., & Tang, J. (2000). Univariate Xbar control charts for individual chractersitcis in a multinormal model. *IIE Transactions Vol 32*, 1115–1125.

Varghese, V., & Rossetti, M. (2008). A Parametric Bootstrapping Approach to Forecast Intermittent Demand. *Proceedings of the 2008 Industrial Engineering Research Conference.*

Williams, T. M. (1982). Reorder Levels for Lumpy Demand. *Journal of the Operational Research Society* , 185–189.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California